EUR-18827
FZKA-6313
ISBN 92-894-2083-9

Probabilistic Accident Consequence Uncertainty Assessment Using COSYMA:

Methodology and Processing Techniques

Prepared by:

J A Jones	National Radiological Protection Board	UK
J Ehrhardt	Forschungszentrum Karlsruhe GmbH	Germany
L H J Goossens	Technical University of Delft	The Netherlands
J Brown	National Radiological Protection Board	UK
R M Cooke	Technical University of Delft	The Netherlands
F Fischer	Forschungszentrum Karlsruhe GmbH	Germany
I Hasemann	Forschungszentrum Karlsruhe GmbH	Germany
B C P Kraan	Delft University of Technology	The Netherlands

CONTENTS

FOREWORD

This is one of a series of reports describing an uncertainty analysis on the predictions of the accident consequence assessment code COSYMA. A complete list of the reports produced in this project is given in Appendix A, where the reports are divided into those describing the expert judgement study on the distributions of the input parameter values and those describing the results of the analysis. This report describes the methodology used in the analysis.

Sections 1.1 and 1.2 of this report are very similar to the same sections in the reports describing the module analyses. The last paragraph of section 1.2 is not included in the module analysis reports. Theses sections are also similar to the equivalent parts of the report on the overall analysis.

Appendices A (list of reports from the project) and B (description of the models in COSYMA) are included in each of the reports on the uncertainty analysis.

ABSTRACT

A study to perform an uncertainty analysis of the European accident consequence assessment system, COSYMA, has been carried out under contract to the European Commission. The study involved a series of analyses of the uncertainty in different sections of the system, followed by a final analysis of the uncertainty in the whole system.

The overall aims of the study can be summarised as:

1 to formulate a state-of-the-art expert judgement methodology which is capable of finding broad acceptance,
2 to apply the methodology to estimate uncertainties associated with the predictions of the probabilistic accident consequence assessment system COSYMA
3 to provide an input to identifying future R&D priorities.

Uncertainty analysis involves specifying probability distributions for the values of each of the parameters involved, sampling sets of values from those distributions and propagating them through the model to derive information on the uncertainty in the model prediction. Those parameters whose uncertainties make major contributions to the overall uncertainty can then be identified using correlation coefficients between the input values and the model outputs.

The study evaluated the uncertainty on air and ground concentration, individual doses and risks, the extent of countermeasures and the numbers of health effects in the population. The calculations were undertaken for a number of situations with and without allowing for the effects of countermeasures. Some licensing procedures require estimates of the potential individual doses and risks at points near the reactor site. Potential doses are calculated assuming people are outdoors for the whole of the period of interest, and so make no allowance for countermeasures or shielding by normal occupation of buildings. The study evaluated such potential doses, and the associated risks of health effects. Other applications of probabilistic risk assessment codes assume that countermeasures will be taken if doses are above selected levels. Such calculations were also considered in this study. Consequences assuming normal living (ie allowing for shielding by buildings but no countermeasures) are considered in the licensing procedures of several countries. Hence calculations were also undertaken for individual and collective doses and risks for normal living.

The source terms chosen encompass a wide range of characteristics (eg magnitude and composition) of source terms that have been postulated for LWRs. They are taken from analyses of the pressurised water reactor proposed for the Hinkley Point site in the UK. UK1 is a very large release; it is the risk-dominant source term for early health effects and a major contributor to the overall risk of late health effects from the reactor. CB2 is a smaller, but less unlikely, sequence that also makes a major contribution to the overall risk of late health effects from the reactor. DBA is a design basis accident.

This report describes the methods used to undertake the analysis, the conditions considered in the study and the reasons for the choices made in planning the study. It gives a summary of the methods used for eliciting expert opinion, combining the distributions obtained from different experts, and obtaining distributions on the input parameters of COSYMA from the distributions obtained from the experts. It describes the steps involved in carrying out the analysis, and justifies the method used to select the uncertain parameters for the final analysis.

ACKNOWLEDGEMENTS

The project team wishes to thank all those people who took part in the expert judgement elicitation process, as members or organisers of the expert panels. The project team also acknowledges a number of useful comments when papers describing the progress on this project have been presented at conferences.

The expert elicitation work and the derivation of distributions on the COSYMA input parameters was partially funded by the European Commission under contract number FI3P-CT92-0023. The module and overall uncertainty analyses were partially funded by the European Commission under contract number FI4P-CT95-0006.

1 BACKGROUND TO THE STUDY

1.1 Introduction

Despite the elaborate precautions taken in the design, construction and operation of nuclear facilities, there will always remain the possibility, however small, of accidental releases of radioactivity into the environment. There is a need to evaluate the risks arising from potential accidents, on a probabilistic basis, taking into account the spectrum of possible consequences of accidents and their associated probability of occurrence. Probabilistic risk assessment (PRA) or accident consequence assessment (ACA) is the process whereby the consequences of potential accidental releases are assessed, taking into account the range of conditions which may prevail at the time of the accident, and the associated probability of these conditions. Such assessments have applications in the design, siting, licensing and operating phases of a nuclear installation. They can be used to evaluate the risks posed by a specific or representative nuclear site, for example for comparison with safety criteria. They can be used for evaluating the effects of design changes or of plant modifications. They also have an input into emergency planning and to some aspects of siting studies.

A number of computer systems have been developed for use in such assessments. Such systems include models for describing the pathways by which people are irradiated following discharges of material, and for calculating the doses and the associated health risks. The models require values to be specified for a large number of input parameters. The predictions of such models are uncertain for two main reasons, which can be summarised as:

(a) *modelling uncertainties,* arising from a lack of knowledge about the most appropriate mathematical formulation to represent environmental processes,
(b) *parameter value uncertainties,* arising from inadequate knowledge about the most appropriate values to be assigned to the many parameters in the model.

The models adopted are not perfect as they contain idealisations and simplifying assumptions. They may not describe all features concerned; features which have been omitted because they make only a small contribution to the "best estimate" model prediction may make larger contributions to the uncertainty. The most appropriate values to be assigned to the many parameters involved in the model may not be known with certainty, leading to uncertainty in the final predictions of the model.

Two computer systems for use in probabilistic accident consequence assessments (COSYMA[1] in the European Union and MACCS[2] in the US) were developed around 1990, and made generally available. There has been an interest in quantifying the uncertainty in the predictions of such systems, and extensive analyses of the uncertainty on predecessors of both programs have been carried out [3,4,5]. An important feature of an uncertainty analysis is the derivation of a joint distribution* on the values of the many parameters involved. In the earlier

* The joint distribution assigns a probability to each feasible set of values of the input parameters.

studies, the joint distribution was largely specified by the system developers, rather than experts in the many different fields involved in accident consequence modelling.

In 1991, both the European Commission (EC) and the United States Nuclear Regulatory Commission (USNRC) were considering initiating studies to better quantify the uncertainty in the input parameter values and in the predictions of the systems. An essential aspect of these studies was to obtain distributions and information on the dependencies between parameter values using formal expert judgement elicitation techniques. The studies were combined into a single EC/USNRC project intended to develop credible and traceable uncertainty distributions for the respective system input parameters. A further intention was for these distributions to be propagated through the two systems, and so quantify the uncertainty in the predictions.

The broad objectives of both the EC and USNRC for this study can be summarised as

1 to formulate a state-of-the-art expert judgement methodology which is capable of finding broad acceptance;

2 to apply the methodology to estimate uncertainties associated with the predictions of the probabilistic accident consequence systems COSYMA and MACCS;

3 to provide an input to identifying future R&D priorities.

Within these broad objectives, small differences in emphasis exist between the EC and USNRC. This report concentrates on the analysis using COSYMA, and the EC aims and objectives.

The first objective was met in two ways. First, the collaboration between research teams from the US and Europe led to the development of agreed methods for the study, and in particular for the formal elicitation of expert judgement. Second, a protocol document describing the methods to be used for the final uncertainty analyses on COSYMA was distributed to a number of researchers in the field for comment. The views expressed on that document have been incorporated into the methods used for the analysis.

The second objective was met by using the joint distribution on the uncertain parameter values derived from the expert elicitation in an analysis of the uncertainty in the predictions of the consequences of accidental releases using COSYMA. Undertaking rigorous uncertainty analyses involves considerable computational costs and substantial effort. It is not possible to carry out such analyses on every occasion when accident consequence assessments are undertaken. It was intended that the levels of uncertainty obtained in this study would indicate the likely levels of uncertainty in other, similar, situations. Therefore, this analysis has been undertaken for several combinations of source term and types of population behaviour with the intention of deriving indicative levels of uncertainty should COSYMA be applied in other situations. For example, if the study shows that the uncertainty in a particular endpoint for a particular countermeasures strategy is a factor of 10, then it can be assumed that in similar situations the uncertainty is also a factor of 10, not 100.

There are several aspects to the third objective above. The uncertainty was better quantified because the distributions on the parameter values are determined from formal techniques of expert judgement. In addition to calculating the uncertainty on the model predictions, the study has also identified the input parameters whose uncertainties make major contributions to the overall uncertainty. This will form an input into identifying research priorities.

Uncertainty analyses can be considered to consist of three broad stages, each of which could be further divided into smaller steps. The first step is to determine what types of uncertainty are present in the model being analysed, which types will be considered in the analysis and which of the model's input parameters will be considered to be uncertain. This step also includes identifying those model endpoints for which the uncertainty will be analysed. The second broad step is to determine the joint distribution on the values of the model input parameters that are being considered. This joint distribution includes not only the ranges of each of the parameter values, but also the probability distribution of the input parameter taking different values within that range and any dependencies between the values of the different parameters within their ranges. In this study, the joint distribution over the model input parameters has been obtained using formal techniques for eliciting expert judgement. These parts of the study have been described in a series of reports, as listed in Appendix A. The final broad step is to sample sets of input parameter values from the joint distribution, to propagate those values through the model, to determine the uncertainty on the model endpoints and identify those parameters whose uncertainties make large contributions to the overall uncertainty.

The models included in COSYMA are described in Appendix B. There are many hundreds of parameters involved in describing the transfer of radioactive material from its release through the environment to man and calculating the subsequent doses and risks. It would not be possible to consider all these parameters in a single analysis, because of the complexity of the analyses and amount of computation that would be required. Therefore, a series of analyses of parts of the complete COSYMA system have been carried out. These are described as "module analyses", although the parts of the code considered in these analyses do not necessarily correspond exactly to the defined modules of COSYMA[1]. Each module includes a number of different models. Those parameters whose uncertainties make major contributions to the overall uncertainty for each module were identified and included in a final overall analysis[6]. The following module analyses were carried out before the final analysis:

1) Dispersion and deposition
2) Foodchain transfer
3) Dosimetry - external, inhalation and ingestion doses
4) Early and late health effects.

These analyses in this study calculated the uncertainty on the overall endpoints of COSYMA coming from the uncertainty in the input parameters for the particular module, rather than simply considering the uncertainty on the endpoints of that particular module. In this way, the importance of the parameter uncertainties can be judged in terms of their contribution to the overall uncertainty and not simply in terms of their contribution to some intermediate quantity in the

calculation. Default values were allocated to the parameters of the other modules for which the uncertainty was not considered in the particular analysis. Thus the analysis of the uncertainty on the dispersion and deposition module assumed default values for the parameters describing food chain transfer, dose models and health effects models. This division into modules is such that no single parameter is input to more than one module, and there are no large correlations between the values of the input parameters for the different modules. This means that the likely contribution of each of the parameters to the overall uncertainty on the COSYMA predictions could be determined from the results of a single module analysis.

Since the study was intended to derive indicative levels for the uncertainty to be expected under normal applications of COSYMA, it was necessary to make as few changes as possible to COSYMA for this analysis. For this reason, the models used in COSYMA were not modified to give a better fit to the distributions provided by the experts. In some cases, the models included in COSYMA are complex and an uncertainty analysis of the full version of the system would have required excessive amounts of computer resources. In these cases the models were simplified so that the uncertainty analysis could be carried out more easily. Simplifications were introduced in the calculation of the risk of late health effects, the models for transfer of some radionuclides to animal products, and the model for human metabolism of actinides. These simplifications will not have significantly altered the extent of the uncertainty on the predictions of COSYMA, though they may have altered slightly the central values about which the uncertainty is expressed. They have not affected the aims of the study, as the objective was to evaluate the extent of the uncertainty in the predictions for typical COSYMA calculations, rather than the absolute value of the consequences of particular accidental releases.

This is one of a series of reports describing the overall analysis of the uncertainty in the predictions of COSYMA. The starting point for this series of reports is taken as the end of the expert elicitation process. Appendix A gives a complete list of the reports relating to the project. This report summarises the methods used in the analysis, including those to determine uncertainty distributions on the COSYMA input parameters from the information provided by the experts. The remainder of this section describes the situations and endpoints for which the uncertainty was evaluated. Section 2 describes the different sources of uncertainty which are considered in the analysis. Section 3 outlines the expert panels used, and the general methods of deriving the information required for the analysis from the information provided by the experts. More detailed information is provided in the reports for each of the module analyses. Section 4 describes the quantities used to describe the uncertainty and gives the reasons behind the choice of method to identify the important parameters from the module analyses for inclusion in the overall analysis.

1.2 Situations considered

Three source terms, encompassing a wide range of characteristics of source terms that have been postulated for LWRs (e.g. magnitude and composition), have been considered in this study. They were taken from analyses of the pressurised water reactor proposed for the Hinkley Point site in the UK. UK1 is a very large release; it was identified as the risk-dominant source term for early health effects and a major contributor to the overall risk of late health effects from the reactor[7]. CB2 is a smaller, but less unlikely, sequence that also makes a major contribution to the overall

risk of late health effects from the reactor[8]. DBA is a design basis accident[9]. This is a fault which the plant is designed to take or can be shown to withstand without unacceptable consequences, by virtue of the plant's inherent characteristics or safety systems. The amounts of material released for the UK1 and CB2 source terms were calculated from the reactor inventory and the release fractions which apply to groups of elements; the amount of each isotope released for the DBA source term was specified directly. The source terms are summarised in Table 1.1 to Table 1.3. Table 1.1 shows the assumed inventory of the reactor;

Table 1.2 gives the release fractions used for the UK1 and CB2 source terms, and Table 1.3 gives the amount of each nuclide released in the DBA source term.

Table 1.2 also gives approximate release fractions for the DBA source term, to enable easy comparisons of the magnitude of this and the other source terms.

The calculations were undertaken for a range of patterns of population behaviour. Some licensing procedures require estimates of the potential individual doses and risks at points near the reactor site. Potential doses are calculated assuming people are outdoors for the whole of the period of interest, and so make no allowance for countermeasures or shielding by normal occupation of buildings. The study evaluated such potential doses, and the associated risks of health effects. Consequences assuming normal living (i.e. allowing for shielding by buildings but no countermeasures) are considered in the licensing procedures of several countries. Hence calculations were also undertaken for individual and collective doses and risks for normal living.

There is also an interest in calculating the uncertainty on the predictions of COSYMA if allowance is made for the countermeasures that might be imposed following a reactor accident. International organisations have suggested ranges of criteria for implementing countermeasures, recognising that intervention levels might depend on the situation and scale of accident that occurs. A countermeasures strategy based on the IAEA[10] intervention levels for sheltering, evacuation, iodine tablets and relocation together with the EU levels for banning food[11,12,13] was used. The intervention levels and implementation times used for this study are given in Table 1.4 Doses and risks are calculated assuming normal living for those not subject to countermeasures, or not subject to countermeasures in a given time period.

COSYMA gives information on a wide variety of consequences of an accident. It was not possible to generate information on all of these endpoints in this study. Therefore, the study evaluated the uncertainty on a selection of endpoints; information on the uncertainty in other endpoints can be deduced from these results. A complete list of endpoints is given in Table 1.5; they can be summarised as follows:

- air concentration and deposition of ^{131}I and ^{137}Cs at selected distances.
- individual dose to 7 days in bone marrow, thyroid and skin at selected distances.
- individual and collective risks of early health effects (total risks of mortality, and of the haematopoietic syndrome, the total risks of morbidities and of lung morbidity and hypothyroidism).
- the areas with emergency actions for sheltering, evacuation and distribution of stable

iodine tablets.

- individual and collective committed effective dose and doses in bone marrow and thyroid.
- individual and collective risks of the numbers of fatal cancers (total and from thyroid) and leukaemia.
- the areas and their time integrals affected by relocation and by food restrictions, for meat, milk, green vegetables and grain.

Different sub-sets of the complete list of endpoints are considered in the different module analyses, as some of the input parameter values for some of the modules do not influence all the endpoints.

The collective health effects were evaluated for a hypothetical site in central Europe, as defined in a recent international intercomparison of reactor accident programs[14].

As stated earlier, the aim of the exercise was to derive indicative levels of uncertainty that should be appropriate for other, similar analyses using COSYMA. The size of uncertainty associated with the predictions may change for different magnitudes of the source term, and for calculations with and without countermeasures. The following set of situations was chosen for analysis, where NE and NL refer to the separate sub-systems of COSYMA relating to the calculation of early effects (NE sub-system) and late effects (NL sub-system):-

UK1 potential outdoor doses and risks, for those NE endpoints relating to individual doses and risks.

UK1 normal living with no countermeasures, for those NE endpoints relating to individual doses and risks, and to numbers of health effects.

UK1 with countermeasures, for those NE endpoints relating to individual doses and risks, and to numbers of health effects.

CB2 normal living with no countermeasures, for those NL endpoints relating to individual doses and risks, collective doses and numbers of late health effects.

CB2 with countermeasures, for all NE and NL endpoints.

DBA potential outdoor doses and risks, for those NL endpoints relating to individual doses and risks.

DBA with countermeasures, for all NL endpoints.

The uncertainty on individual doses and risks for early effects (the NE endpoints) were evaluated at 0.875, 5 and 20 km, while the uncertainties on individual doses and risks for late effects (the NL endpoints) were evaluated at 5, 20 and 100 km. COSYMA calculates doses at discrete points on a spatial grid, and assumes that the dose at the centre of each grid area applies throughout that area. Thus the dose at 0.875 km is calculated as representing the doses over the distance band between 0.75 and 1 km.

This combination of conditions means that information on the uncertainty of the numbers of early health effects in the population was obtained mainly from the analyses for the UK1 source

term. Little information on the uncertainty on these endpoints could be obtained from the analyses with the CB2 source term as doses from this source term were generally below the thresholds for producing early health effects. Information on the uncertainties in doses over short time periods and risks of early health effects for people who are outdoors at the time of the accident, for people who are living normally with no countermeasures taken, and if countermeasures are taken on the basis of doses in the exposed population were obtained from the analyses for the UK1 source term. The predicted risks of early health effects, and the associated uncertainties in the predictions, will not depend on the criteria used to invoke countermeasures unless they are such that some people who receive doses above the threshold for deterministic effects are not sheltered and evacuated. Although the analysis for the CB2 source term could not give much information on risks of early health effects, it did give results for the doses in short time periods, both for normal living and if countermeasures were taken.

Information on the uncertainty in the predicted extent of early countermeasures (sheltering, evacuation and distribution of stable iodine tablets) was obtained from the analyses for the CB2 source term. Information on the uncertainty on the late countermeasures (relocation and food restrictions) was obtained from the analyses for the CB2 and DBA source terms. Two source terms were selected for this part of the analysis as they have different relative contributions from the iodine and caesium isotopes.

Information on the predicted risks of late health effects was also obtained from the CB2 and DBA source terms, for both individual and collective risks. Again, the two source terms were used because of the different relative contributions of the iodine and caesium isotopes.

The extent of the uncertainty on the predicted air concentration and deposition does not depend on the size of the release. The endpoints relating to concentration and deposition were only considered in the analysis for the CB2 source term, as this is the only source term for which all four distances (from NE and NL) were considered.

The results from a single run of COSYMA are presented using the complementary cumulative frequency distribution function (ccdf), which gives the probability that the consequence is greater than a particular value. The distribution can be summarised using various characteristic quantities such as the probability of zero effects, the expectation value (the mean or average of the distribution) and various percentiles. The *nth* percentile is the level of consequence that is exceeded with a probability of (100-n) percent. The quantities used to represent the uncertainty are described in Section 4 of this report.

The original intention was to present results for the uncertainty on the mean value, the 95^{th} and 99^{th} percentiles of the ccdf and the probability of zero consequences. Unfortunately, difficulties were experienced in analysing the results for the probability of zero. These partly reflect the way in which COSYMA calculates the ccdfs; COSYMA calculates the probability that the consequence is lower than a very small value rather than the true probability of zero. This lead to difficulties in places where the reported probability of zero was uncertain in situations where the module parameters could not influence the probability of a true zero consequence. The project staff therefore decided, at a late stage in the analysis, that results for the uncertainty on the probability of

zero would not be presented. This decision was taken after the parameters to be included in the overall analysis had been selected. Consequently the overall analysis includes some parameters that are only important because of their influence on the uncertainty of the probability of zero. This does not affect the results of the study as the overall analysis included more, rather than less, parameters than would have been included if the change had been made earlier.

1.3 References

1. KfK and NRPB. COSYMA - a new program package for accident consequence assessment. CEC Brussels, EUR 13028 (1991)

2. Chanin D I et al. MELCOR accident consequence code system (MACCS), User's Guide. NUREG/CR-4691. Albuquerque (1990)

3. Jones J A, Mansfield P A and Crick M J. An uncertainty analysis of the predicted consequences of nuclear accidents using the NRPB code MARC-2A. Chilton, NRPB-R274 (London HMSO) (1995)
4. Fischer F, Ehrhardt J and Hasemann I. Uncertainty and sensitivity analyses of the complete program system UFOMOD and of selected submodels. Karlsruhe KfK 4627 (1990)
5. Kocher D C, Ward R C, Killough G G et al. Sensitivity and uncertaitny studies of the CRAC2 computer code. NUREG/CR-4038 (1985)
6. J A Jones, J Ehrhardt, L H J. Goossens, F Fischer, I Hasemann, B C P Kraan, R M Cooke. Probabilistic accident consequence uncertainty assessment using COSYMA: Overall uncertainty analysis. EUR 18826 and FZKA 6312 (2000)
7. Jones J A and J A Williams J A. An assessment of the radiological consequences of releases from degraded core accidents from a proposed PWR at Hinkley Point: Results using MARC-1. Chilton, NRPB-M152 (1988)
8. Jones J A and Williams J A. An assessment of the radiological consequences of releases from containment bypass accidents from a proposed PWR at Hinkley Point: Results using MARC-1. Chilton. NRPB-M154 (1988)
9. Jones J A and Williams J A. An assessment of the radiological consequences of releases from design basis accidents from a proposed PWR at Hinkley Point: Results using MARC-1. Chilton. NRPB-M153 (1988)
10. FAO, IAEA, ILO, OECD/NEA, PAHO, WHO. International basic safety standards for protection against ionizing radiation and for the safety of radiation sources. Vienna, IAEA, Safety Series 115 (1996)
11. CEC. Council Regulation (Euratom) No 3954/87 laying down the maximum permitted levels of radioactive contamination of foodstuffs and feedingstuffs following a nuclear accident or any other case of radiological emergency. *Off. J. Eur. Commun.* L371/1/11 (1987), amended by Council Regulation 2218/89. *Off. J. Eur. Commun.* L211/1 (1989).
12. CEC. Council Regulation (Euratom) No 944/89 laying down the maximum permitted levels of radioactive contamination in minor foodstuffs following a nuclear accident or any other case of radiological emergency. *Off. J. Eur. Commun.* L101/17 1989.
13. CEC. Council Regulation (Euratom) No 770/90 laying down maximum permitted levels of radioactive contamination of feedingstuffs following a nuclear accident or any other case of

radiological emergency. *Off. J. Eur. Commun.* L83/78 (1990).

14. Nuclear Energy Agency and Commission of the European Communities. Probabilistic accident consequence assessment codes. Second international comparison. Paris, OECD (1994)

Table 1.1 Reactor inventory considered

Radionuclide	Inventory (Bq)	Half-life	Radionuclide	Inventory (Bq)	Half-life
^{58}Co	3.08 10^{16}	70.8 d	^{131m}Te	3.47 10^{17}	30.0 h
^{60}Co	1.14 10^{16}	5.27 y	^{132}Te	4.85 10^{18}	78.2 h
^{85}Kr	2.17 10^{16}	10.7 y	^{131}I	3.39 10^{18}	8.04 d
^{85m}Kr	9.25 10^{17}	4.48 h	^{132}I	4.96 10^{8}	2.30 h
^{87}Kr	1.70 10^{18}	76.3 min	^{133}I	6.81 10^{18}	20.8 h
^{88}Kr	2.34 10^{18}	2.84 h	^{134}I	7.84 10^{18}	52.6 min
^{86}Rb	7.96 10^{15}	18.6 d	^{135}I	6.40 10^{18}	6.61 h
^{89}Sr	3.37 10^{18}	50.5 d	^{133}Xe	6.85 10^{18}	5.25 d
^{90}Sr	1.75 10^{17}	29.1 y	^{135}Xe	1.67 10^{18}	9.09 h
^{91}Sr	4.37 10^{18}	8.48 h	^{134}Cs	3.85 10^{17}	2.06 y
^{90}Y	1.82 10^{17}	2.67 d	^{136}Cs	1.33 10^{17}	13.2 d
^{91}Y	4.51 10^{18}	58.6 d	^{137}Cs	2.29 10^{17}	30.0 y
^{95}Zr	5.88 10^{18}	65.5 d	^{140}Ba	6.14 10^{18}	12.7 d
^{95}Nb	5.81 10^{18}	35.1 d	^{140}La	6.32 10^{18}	40.3 h
^{97}Zr	5.88 10^{18}	16.9 h	^{141}Ce	5.92 10^{18}	32.5 d
^{99}Mo	6.44 10^{18}	66.02 h	^{143}Ce	5.44 10^{18}	33.0 h
^{99m}Tc	5.55 10^{18}	6.02 h	^{144}Ce	3.59 10^{18}	285 d
^{103}Ru	5.25 10^{18}	39.4 d	^{143}Pr	5.40 10^{18}	13.6 d
^{105}Ru	3.51 10^{18}	4.44 h	^{147}Nd	2.36 10^{18}	11.0 d
^{106}Rh	3.18 10^{18}	1.47 d	^{239}Np	7.32 10^{19}	2.36 d
^{106}Ru	1.30 10^{18}	368 d	^{238}Pu	3.17 10^{15}	87.7 y
^{127}Sb	2.93 10^{17}	3.89 d	^{239}Pu	1.11 10^{15}	2.41 10^{4} y
^{129}Sb	9.95 10^{17}	4.31 h	^{240}Pu	1.06 10^{15}	6550 y
^{127}Te	2.85 10^{17}	9.35 h	^{241}Pu	3.12 10^{17}	14.4 y
^{127m}Te	4.37 10^{16}	109 d	^{241}Am	2.06 10^{14}	432 y
^{129}Te	9.40 10^{17}	69.6 min	^{242}Cm	6.62 10^{16}	163 d
^{129m}Te	1.67 10^{17}	33.6 d	^{244}Cm	2.75 10^{15}	18.1 y

Table 1.2 Source terms considered for the assessment

Source term	Fraction of core inventory released to the environment								
	Xe-Kr	Organic iodine	Inorganic iodine	Cs-Rb	Te-Sb	Ba-Sr	Ru[(a)]	La[(b)]	Pu[(c)]
UK1	9 10^{-1}	7 10^{-3}	7 10^{-1}	5 10^{-1}	3 10^{-1}	6 10^{-2}	2 10^{-2}	4 10^{-3}	4 10^{-3}
CB2	1 10^{-2}	5 10^{-6}	2 10^{-3}	8 10^{-3}	8 10^{-6}	8 10^{-7}	8 10^{-7}	8 10^{-7}	3 10^{-7}
DBA[(d)]	1 10^{-7}	-	1 10^{-6}	1 10^{-6}	1 10^{-8}	1 10^{-8}	1 10^{-8}	1 10^{-8}	1 10^{-10}

Notes

a Includes Ru, Rh, Co, Mo, Tc.
b Includes Y, La, Zr, Nb, Ce, Pr, Nd.
c Includes Np, Pu, Am, Cm.
d This source term is defined in terms of the amount of each radionuclide released. The information has been converted into the form presented here for comparison with the other source terms. The release fractions for different isotopes of the same element and for different elements differ from the values given here by up to a factor of 3.

Table 1.3 Activity released in the DBA source term

Radionuclide	Release (Bq)	Radionuclide	Release (Bq)	Radionuclide	Release (Bq)
^{24}Na	$7.0\ 10^{10}$	^{51}Cr	$1.4\ 10^{11}$	^{54}Mn	$1.4\ 10^{11}$
^{55}Fe	$5.2\ 10^{9}$	^{59}Fe	$5.2\ 10^{9}$	^{58}Co	$3.4\ 10^{11}$
^{60}Co	$3.2\ 10^{10}$	^{63}Ni	$5.6\ 10^{9}$	^{65}Zn	$1.4\ 10^{11}$
^{83}Br	$9.3\ 10^{10}$	^{84}Br	$2.6\ 10^{12}$	$^{85}Br^{(a)}$	$4.8\ 10^{9}$
^{83m}Kr	$5.2\ 10^{9}$	^{85m}Kr	$1.1\ 10^{11}$	^{85}Kr	$2.3\ 10^{9}$
^{87}Kr	$9.3\ 10^{10}$	^{88}Kr	$1.1\ 10^{11}$	^{89}Kr	$8.1\ 10^{10}$
^{86}Rb	$4.4\ 10^{9}$	^{88}Rb	$3.5\ 10^{13}$	^{89}Rb	$8.1\ 10^{12}$
^{89}Sr	$4.4\ 10^{10}$	^{90}Sr	$3.7\ 10^{8}$	^{91}Sr	$2.3\ 10^{11}$
^{90}Y	$4.4\ 10^{8}$	^{91m}Y	$6.3\ 10^{10}$	^{91}Y	$4.8\ 10^{8}$
^{93}Y	$3.7\ 10^{11}$	^{95}Zr	$4.1\ 10^{10}$	^{95}Nb	$4.4\ 10^{10}$
^{99}Mo	$1.6\ 10^{11}$	^{99m}Tc	$3.7\ 10^{10}$	^{103}Ru	$2.7\ 10^{10}$
^{106}Ru	$1.6\ 10^{10}$	^{103m}Rh	$6.3\ 10^{10}$	^{106}Rh	$3.5\ 10^{10}$
^{110m}Ag	$5.6\ 10^{10}$	^{122}Sb	$1.0\ 10^{11}$	^{124}Sb	$2.5\ 10^{10}$
^{125m}Te	1.7 10	^{127m}Te	$1.8\ 10^{9}$	^{127}Te	$8.5\ 10^{9}$
^{129m}Te	$3.3\ 10^{10}$	^{129}Te	$8.9\ 10^{12}$	^{131m}Te	$1.2\ 10^{11}$
^{131}Te	$2.3\ 10^{12}$	^{132}Te	$1.8\ 10^{10}$	^{130}I	$1.9\ 10^{10}$
^{131}I	$1.9\ 10^{12}$	^{132}I	$5.2\ 10^{12}$	^{133}I	$8.1\ 10^{12}$
^{134}I	$6.3\ 10^{12}$	^{135}I	$3.6\ 10^{12}$	^{131m}Xe	$2.3\ 10^{10}$
^{133m}Xe	$2.8\ 10^{10}$	^{133}Xe	$1.5\ 10^{12}$	^{135m}Xe	$9.3\ 10^{10}$
^{135}Xe	$3.4\ 10^{11}$	^{137}Xe	$8.1\ 10^{11}$	^{138}Xe	$4.1\ 10^{11}$
^{134}Cs	$2.1\ 10^{11}$	^{136}Cs	$2.5\ 10^{10}$	^{137}Cs	$2.7\ 10^{11}$
^{138}Cs	$5.9\ 10^{12}$	^{139}Cs	$2.0\ 10^{13}$	^{137m}Ba	$8.9\ 10^{11}$
^{139}Ba	$4.4\ 10^{12}$	^{140}Ba	$6.7\ 10^{10}$	^{140}La	$3.5\ 10^{10}$
^{141}Ce	$1.0\ 10^{10}$	^{143}Ce	$3.7\ 10^{10}$	^{144}Ce	$3.7\ 10^{10}$
^{143}Pr	$3.6\ 10^{8}$	^{144}Pr	$3.7\ 10^{10}$	^{187}W	$2.2\ 10^{11}$
^{237}U	$2.5\ 10^{8}$	^{239}U	$1.0\ 10^{10}$	^{239}Np	$4.1\ 10^{9}$
^{236}Pu	$1.7\ 10^{5}$	^{238}Pu	$3.7\ 10^{5}$	^{239}Pu	$1.5\ 10^{5}$
^{240}Pu	$1.4\ 10^{5}$	^{241}Pu	$4.1\ 10^{7}$	^{242}Pu	$4.4\ 10^{2}$
^{243}Pu	$8.5\ 10^{7}$	^{241}Am	$7.0\ 10^{4}$	^{242m}Am	$2.4\ 10^{3}$
^{242}Am	$4.8\ 10^{7}$	^{243}Am	$8.1\ 10^{3}$	^{244}Am	$2.7\ 10^{6}$
^{242}Cm	$1.6\ 10^{6}$	^{243}Cm	$6.3\ 10^{2}$	^{244}Cm	$9.6\ 10^{4}$

Table 1.4 Countermeasures criteria and timings adopted in the study

Action	**Criteria**		
Sheltering	10 mSv effective dose, total of committed inhalation dose and external dose in 7 days to a person outdoors		
Evacuation	50 mSv effective dose, total of committed inhalation dose and external dose in 7 days to a person outdoors		
Iodine tablets	100 mSv inhalation dose to thyroid to a person outdoors		
Relocation	30 mSv external dose in 30 days for normal living		
Return from relocation	10 mSv external dose in 30 days for normal living		
Food restrictions	Activity concentration levels in food		
	Radionuclide	Milk (Bq l^{-1})	Other foods (Bq kg^{-1})
	Strontium	125	750
	Iodine	500	2000
	Caesium and other long-lived radionuclides	1000	1250
	α - emitters	20	80

Action	**Time when action initiated**	**Time when action withdrawn**
Sheltering	2 hours	8 hours
Evacuation	6 hours	2 days
Iodine tablets	4 hours	-[a]
Relocation	Depends on relocation area[b]	When dose rate drops below criterion
Food restrictions	Start of first time period in which concentrations are above the criterion	End of last time period in which concentrations are above the criterion

Notes:

a COSYMA assumes that iodine tablets are taken on a single occasion only.
b COSYMA calculates an average relocation time, assuming that the area affected can be relocated at a rate of 100 km^2 per day, and assumes that everyone is relocated at that time

Table 1.5 List of endpoints considered in the analysis

For COSYMA NE[a] runs

Activity concentrations, at 0.875, 5 and 20 km.
in air and on the ground, for Cs-137 and I-131.

Individual doses, at 0.875, 5 and 20 km
integrated to 7 days for both inhalation and external dose
for bone marrow, thyroid and skin.

Individual risks of deterministic health effects, at 0.875, 5 and 20 km.
for mortality, the sum and the risk of the haematopoietic syndrome,
for morbidity, the sum and the risk of lung morbidity, hypothyroidism and skin burns.

Areas with emergency actions,
for sheltering only, evacuation and distribution of stable iodine tablets.

Number of deterministic health effects
for mortality, the sum and haematopoietic syndrome.
for morbidity, the sum and numbers of cases of lung morbidity, hypothyroidism and of skin burns.

For COSYMA NL[b] runs

Activity concentrations, at 5, 20 and 100 km
in air and on the ground, for Cs-137 and I-131.

Individual doses, at 5, 20 and 100 km
integrated to 50 years for both inhalation and external dose
effective dose and for bone marrow and thyroid.

Individual risk of fatal stochastic health effects, at 5, 20 and 100 km
for total, and the risks of death from leukaemia and thyroid cancer.

Areas with countermeasures
for relocation, the initial area and its time integral
for restrictions of milk, grain, leafy vegetables and beef, the initial area and its time integral.

Collective doses
effective dose and for bone marrow and thyroid.

Numbers of fatal stochastic health effects
the sum, and numbers of deaths from leukaemia and thyroid cancer.

Notes:

a: NE refers to the sub-system of COSYMA calculating short term doses, early health effects and the appropriate countermeasures

b: NL refers to the sub-system of COSYMA calculating long term doses, late health effects and the appropriate countermeasures

2 ASPECTS OF THE ANALYSIS

2.5 Types of uncertainty included in the analyses

The predictions of models for probabilistic accident consequence assessments are uncertain for a variety of reasons, as summarised in Section 1. This study considers the uncertainty in the results reflecting the uncertainty in the most appropriate values to allocate to the input parameters. The distributions on the values of the input are specified in a way which includes to some extent the uncertainty in the models used in COSYMA. The uncertainty on the values to be assigned to the input parameters can be considered to fall into two different forms, which have been designated as "type A" and "type B" uncertainties respectively[1]. Type A uncertainties have also been identified as "objective" or "stochastic", while type B uncertainties have also been identified as "subjective" or "cognitive". The term "objective" indicates that there is no single correct value for the parameter; the term "stochastic" indicates that the uncertainty is at least partially the result of random physical processes. The terms "subjective" and "cognitive" indicate that there is, in theory, a correct value but it is not known because of a lack of information about a deterministic process. In practice, the difference between the types of uncertainty is often blurred, with parameters having some aspects of both types. COSYMA draws a clear distinction between the two types of uncertainty, with results from normal runs of the program presented as probability distributions that include the stochastic uncertainty on the atmospheric conditions. Therefore the two types of uncertainty are kept separate in this study.

A full description of the risk of hypothetical accidents must allow for the wide range of possible conditions that could prevail at the time of the accident, and the probability distribution of those conditions. PRA involves calculating the risk of accidental releases in a series of conditions that represent the complete range of conditions that might occur at the time of an accident. The probability distribution of consequences is then generated by allocating a probability to the different conditions that can occur. In COSYMA, the atmospheric conditions at the time of the accident are the only aspect for which a probabilistic treatment has been adopted in the uncertainty analysis; other processes which could be modelled in a probabilistic manner, such as the behaviour of the population in the event of emergency actions and the distribution of driving times out of the affected area, have been approximated in COSYMA by a single value allocated to the corresponding parameters.

The "stochastic" uncertainties are addressed in a smaller analysis of the uncertainty arising from the meteorological sampling scheme, by comparing the results for different choices of atmospheric sequences for analysis. The "subjective" uncertainties are addressed in the module and overall analyses by selecting values for the important parameters from distributions describing the possible ranges of those values.

The uncertainty in the predicted consequences of an accident in a defined set of conditions could be reduced substantially if sufficient research were carried out that the appropriate value of each input parameter could be assigned with certainty (i.e. cognitive uncertainty is reduced to zero). By contrast, the atmospheric conditions at the time of a future accident, rather than the

probability distribution of possible conditions, can never be specified with certainty (i.e. stochastic uncertainty cannot be reduced)

Probabilistic risk assessments which do not consider cognitive uncertainty, but do take account of stochastic uncertainty, are carried out by selecting a series of atmospheric conditions according to some appropriate sampling scheme, and calculating the consequences should the accident happen in each of the selected conditions. A probability distribution of consequences is then generated by allocating each of the consequences the probability associated with the selected sequence of atmospheric conditions. However, it is impossible to calculate in advance the consequences of the accident in every possible sequence of conditions that might occur over the life-time of a nuclear plant, and so a representative set of conditions is used. A somewhat different probability distribution would be obtained if a different series of sequences of atmospheric conditions had been selected. The predicted consequences of the accident are thus uncertain because of the choice of atmospheric conditions, even if the probability of different groups of atmospheric conditions occurring is known with certainty.

The main parts of this study have been concerned with determining the uncertainty in the predicted consequences of accidental releases reflecting the uncertainty in the most appropriate value to assign to the many parameters in the model. However, a separate study has also been carried out of the uncertainty from different choices of sets of atmospheric conditions to consider in the analysis. The uncertainties from the choice of meteorological sampling scheme are compared with the uncertainties reflecting the choice of parameter values in the final report on the study[2].

As stated in Section 1, predictions of the consequences of accidental releases are uncertain both because the most appropriate form of the model is uncertain and because the most appropriate values for the many input parameters are also uncertain. Formal techniques have been developed to determine the uncertainty on the predictions of a model reflecting the uncertainty in the values of the parameters (i.e. cognitive uncertainty), but there are currently no formal techniques available to quantify the effects of model uncertainty. This study has concentrated on the uncertainty arising from the lack of knowledge of the most appropriate value to assign to the model parameters. A distribution has been assigned to the value of each of the parameters to represent the probability that the parameter might take different values.

The distributions were obtained using formal techniques for eliciting expert judgement, as described in Section 3. In this process, the experts were asked for distributions on observable quantities rather than on more abstract model parameters. Distributions for the input parameters of the models used in COSYMA were then obtained by fitting the model predictions to the distributions obtained from the experts. In this way, some aspects of model uncertainty are included in the distributions derived for the input parameters. The reports on the module analyses describe the uncertainties that were included in the distributions provided by the expert panels. The methods used to derive distributions on the input parameter values from the distributions on observable quantities are summarised in Section 4 and described in more detail in Appendix C.

In addition to considering the uncertainty on the values to be assigned to the different parameters, the study considered the dependencies between the distributions of the different

parameters. The joint distribution among the code input parameters is represented by marginal distributions for each parameter together with a rank correlation matrix. Correlations between the code input parameters were obtained in three ways:

(1) In the case where code input parameters are observable quantities, the experts were queried on the probability that the values of pairs of parameters would both be above their median values, if they could be measured simultaneously. Given certain assumptions, the conditional probabilities can be transformed into correlation coefficients. This process is described in more detail in Section 3 of this report.

(2) When the distributions on the code input parameters were obtained by processing information obtained from the expert panel on other quantities, a joint distribution on related sets of parameter values is naturally obtained. Correlations were extracted from this joint distribution.

(3) The expert panels were only asked about those parameters for which the organisers of the panels considered that their uncertainty would make major contributions to the overall uncertainty. Project staff have specified distributions for parameters not considered by the expert panels, and have also specified the correlations between those values.

Following an accidental release of material, countermeasures would probably be invoked to reduce the impact of the accident on the exposed population. Evaluating the impact of such actions on the consequences of an accident requires values to be specified for the criteria at which the actions would be initiated and withdrawn, and also for the timings of those actions.

The distributions on most of the parameter values represent the degree of belief that a physically observable quantity will take a particular value. This form of uncertainty is different from that involved in determining what criteria would be applied for invoking countermeasures following an accidental release. Here the distributions would reflect the uncertainty on the choices made by the emergency management team. There is presumably no uncertainty in the decision that a particular decision-maker would make following an accident, though different decision-makers would make alternative decisions. This uncertainty cannot be quantified by current methods of expert judgement elicitation. Therefore the uncertainty on the criteria used for invoking countermeasures has not been included in the present analysis.

The uncertainty on some aspects of the timings of countermeasures also cannot be quantified by the current methods for expert judgement elicitation. Some timings could be affected by decisions made by the emergency management team or by the length of time required to monitor the situation. Others, such as the time taken for people to leave the contaminated area once the decision to evacuate has been made, depend on human behaviour. Some of these uncertainties are being derived in a parallel exercise, but they have not been included in this study.[3]

2.6 Choice of sequences of atmospheric conditions for the analyses

ACA codes predict the probability distribution of consequences should an accident occur

in any of the wide range of atmospheric conditions which might occur at the site of interest during the period in which the site operates. The sequences of conditions are obtained by using a data file giving atmospheric conditions every hour over a period of a few years, and assuming that the conditions during the future operation of the site will be similar to those observed in the past. It is impossible to undertake the calculations for every sequence of conditions over the operating period of the site in advance, and even considering every sequence on a data file used would require excessive computer resources. Therefore a representative sample of starting times must be used.

There are three ways in which the starting times can be selected. The simplest systems are random and cyclic sampling. In random sampling the starting times are simply chosen at random from those in the data file. In cyclic sampling, the starting times are taken to be every *nth* hour on the file, starting from a selected point on the file. For random or cyclic sampling, the consequences for the different starting times are given equal probability when constructing the probability distribution of consequences. The third, and more complicated method, is to use stratified sampling. In this system, the starting times are grouped into those likely to give similar predicted consequences, and one or more hours selected at random from each of the groups. The consequences predicted for that hour are then assigned a probability reflecting the fraction of the total hours on the data file falling into the particular group.

There are problems with all of these systems. Random and cyclic sampling tend to select those conditions that occur more frequently, and so may give a poor prediction of the high percentiles of the probability distribution. The main difficulty with stratified sampling is to identify a sampling scheme where the variation between the consequences corresponding to the different sequences allocated to a group is small compared to the differences between the consequences for sequences in different groups. A further difficulty is the possible mis-allocation of a sequence with high predicted consequences, and low correct probability, to a group with much higher probability. This would occur if the high consequences stem from some particular feature which is not considered in setting up the sampling scheme. Stratified sampling schemes could be dependent on the endpoint of interest. For example, the predicted numbers of early health effects and early countermeasures will be determined by sequences of conditions within the first few km of the site, while the predicted numbers of late health effects will be determined by sequences of conditions occurring while the plume travels for several hundred km from the site.

Stratified sampling schemes should consider atmospheric conditions occurring over the distance range in which most of the consequences arise. Conditions occurring when the plume is more than a few km from the site will not affect the predicted numbers of early health effects, and so would not be appropriate as the basis for a sampling scheme for early health effects. Conversely, late health effects can occur at any distance from the site, and so atmospheric conditions throughout the plume's travel could affect the predicted number of late health effects. Therefore it would be appropriate to base a sampling scheme for late health effects on conditions over a wide range of distances from the site.

Stratified sampling schemes have been widely used in analyses with most ACA programs. In particular, schemes have been suggested for use with COSYMA, and also with MARC[4] and UFOMOD[5], the programs on which COSYMA is based. However, the studies

undertaken in deriving these schemes were concerned mainly with calculating early health effects, and very little emphasis has been given to late health effects and long term countermeasures.

Stratified sampling introduces uncertainty into the final predictions for two closely related reasons, namely the choice of the sampling scheme and the selection of representative starting times from within each group of the scheme. Random and cyclic sampling introduce uncertainty into the overall predictions from the choice of starting times. Information on the relative magnitude of these uncertainties can be obtained from two recent studies.

1 Part of the CEC/NEA COSYMA user's benchmark study[6] was a calculation in which participants chose their own meteorological sampling scheme, but where all other parameters were specified. This gives an indication of the uncertainty in the predictions reflecting the difference between sampling schemes. KEMA undertook further calculations comparing consequences calculated in different runs using the same sampling scheme but taking different sequences from each of the groups.

2 Jones[7] presented a paper at the second COSYMA user group meeting giving results from several runs of a set of sampling schemes. Several runs of COSYMA were undertaken for each sampling scheme, taking different sequences from each group.

Both studies showed that the spread of consequences from different sampling schemes is comparable to that when different sequences are obtained from the same sampling scheme.

Jones[7] presented results of calculations of a range of endpoints using different stratified sampling schemes, chosen to be appropriate for different endpoints. All endpoints were calculated using three sets of sequences taken from each of the sampling schemes. This study showed that the ranges of predicted consequences arising from different sampling schemes are comparable to the ranges arising when different sequences are taken from the same scheme. There is little evidence that the spread of predicted consequences when different sets of sequences are taken from one sampling scheme is smaller for the “appropriate” sampling schemes than for the “inappropriate” sampling schemes.

The work on choice of sampling scheme generally suggests that stratified sampling schemes should distinguish firstly between those sequences with and without rain. The dry sequences should then be further divided into groups of stability category and wind speed, and the wet sequences into groups of rainfall rate perhaps also distinguishing the distance from the site at which the rain falls. The work with MARC in particular suggests that the wet sequences should be categorised in terms of the amount of material deposited, rather than simply the occurrence of rain. This reflects the important contribution of dose from deposited material, particularly to early health effects, and the large variation in deposition with changes in rainfall rate. This means that any stratified sampling scheme used in the uncertainty analysis should also be defined in terms of the amount of material deposited. The allocation of starting times to groups is then modified by the value assigned to the washout coefficient for the particular run of COSYMA.

Work with the choice of sampling scheme for COSYMA, using the Benchmark meteorological data which was used for the uncertainty analysis, showed that it is difficult to define

the grouping scheme for wet conditions in such a way that all groups have some sequences, when allowance is made for different wind directions. This would be more difficult for the uncertainty analysis as the value of the washout coefficient selected for each run would affect the allocation of starting hours to groups. This would almost certainly mean that the different runs within the analysis have different numbers of sequences. This may or may not be a desirable feature, but it could complicate the later parts of the analysis when important uncertain parameters are identified.

A further aspect of meteorological sampling was considered in determining the scheme to use in this analysis. The study is intended to analyse the uncertainty on the expectation value and the 95th and 99th percentiles of the probability distributions of consequences. Stratified sampling was originally introduced to improve the calculations of the highest percentiles (eg the 99.9th percentile) of the probability distribution. The choice of sampling scheme is less important for analyses that are not considering such high percentiles.

For these reasons, the main analyses were undertaken using cyclic sampling, with the same set of sequences used for each of the analyses for which full probabilistic runs were used. The uncertainty due to meteorological sampling was analysed in a separate study carried out alongside the main uncertainty analysis, and described in appendices to the report describing the overall uncertainty analysis[2]. The meteorological sampling part of this study considered different methods of cyclic sampling and stratified sampling schemes based on those developed for UFOMOD and COSYMA.

2.7 Choice of nuclides to consider

The nuclear fission process produces many radionuclides in a nuclear reactor. These decay, and in many cases, the daughter products are also radioactive. The processes result in there being several hundred radionuclides in a reactor, and potentially several hundred radionuclides in the source term. The amount of each nuclide released for the UK1 and CB2 source terms was expressed in terms of the fraction of the core inventory of different elements that are released - these fractions are specified in Section 1.

Many of these nuclides have a very short half-life and so make a negligible contribution to the predicted consequences of the accidental release. Calculations of the consequences of the assumed releases which consider the contributions of all nuclides are not feasible because of the computer resources, both in terms of processing time and disk storage, that would be required. COSYMA includes a method of selecting those nuclides in the release that give the largest doses, and only those nuclides are considered in the analysis.

To do this, COSYMA evaluates a quantity that is related to the dose that would be received from each of the nuclides in the source term. For example, the quantity that is considered for the inhalation dose is the product of the dose coefficient and the amount of the nuclide released. The quantity that is considered for the external dose from deposited material is the product of the amount of each nuclide released, its deposition velocity and the dose per unit deposit over an appropriate time period. The contributions are then ranked in order, and the nuclides giving a selected fraction of the dose from the release are selected for consideration. The selection is undertaken separately for each route of exposure considered, and for doses integrated over time

periods appropriate to early and late health effects. The COSYMA calculations are then undertaken for those nuclides that are identified for each route of exposure, so that different nuclides are considered for the different exposure pathways. The nuclides selected for this study are those that give a total of 99% of the calculated dose for each pathway, if the selection is made using the default values of the parameters considered.

The nuclides considered for the different routes of exposure are shown in Table 2.1 for the UK1 source term, in Table 2.2 for the CB2 source term and in Table 2.3 for the DBA source term.

2.8 References

1. IAEA. Evaluating the reliability of predictions made using environmental transfer models. IAEA Safety Series 100. Vienna (1989)

2. Jones J A, Ehrhardt J, Goossens L H J, Fischer F, Hasemann I, Kraan B C P and Cooke R M.. Probabilistic accident consequence uncertainty assessment using COSYMA: Overall uncertainty analysis. EUR 18826 and FZKA 6312 (2000)

3. Goossens L H J, Jones J A, Ehrhardt J, Kraan B C P. Probabilistic accident consequence uncertainty assessment: Countermeasures Uncertainty Assessments. EUR 18821 and FZKA 6307 (2000)

4. Hill M D, Simmonds J R and Jones J A. NRPB methodology for assessing the radiological consequences of accidental releases of radionuclides to atmosphere – MARC-1. Chilton, NRPB-R224 (1988)

5. Ehrhardt J, Burkart K, Hasemann I, Matzerath C, Panitz H-J and Steinhauer C. The program system UFOMOD for assessing the consequences of nuclear accidents. KfK-4330 (1988)

6. van Wonderen E, van der Steen J, Hasemann I. COSYMA: users intercomparison exercise. EUR 15108 (1994)

7. Jones J A. Meteorological sampling in COSYMA. Presented at the COSYMA user group meeting, Budapest June 1995, and in the proceedings, KEMA report 40666-NUC 95-5947 (1995)

Table 2.1 Nuclides included in the UK1 source term

Deposited γ dose

Zr 95	Ru	103, 106	
Te 132	I	131, 132, 133, 134, 135	
Cs	134, 136, 137	Ba	140
La 140			

Cloud γ dose

Kr 85m, 87, 88	Rb	88	
Sr 91	Sb	129	
Te 131m, 132	I	131, 132, 133, 134, 135	
Xe	133, 135	Cs	134, 136, 137
Ba	140	La	140
Np	239		

Inhalation

Rb	88	Sr	89, 90, 91
Y 91	Zr	95, 97	
Nb	95	Mo	99
Ru	103, 105, 106	Rh	105
Sb	127, 129	Te	127, 127m, 129, 129m, 131m, 132
I 131, 132, 133, 134, 135		Cs	134, 136, 137
Ba	140	La	140
Ce	141, 143, 144	Pr	143
Nd	147	Np	239
Pu	238, 239	Cm	242, 244

Skin

Sr 89, 91	Mo	99	
Te 127, 132	I	131, 132, 133, 134, 135	
Cs	134, 136	Ba	140
La 140	Np	239	

Table 2.2 Nuclides included in the CB2 source term

Deposited γ dose

I	131, 132, 133, 134, 135	Cs	134, 136, 137

Cloud γ dose

Kr	85m, 87, 88	Rb	88
I	131, 132, 133, 134, 135	Xe	133, 135
Cs	134, 136		

Skin

I	131, 132, 133, 134, 135	Cs	134, 137

Inhalation

Co	60	Rb	86, 88
Sr	89, 91	Y	91
Zr	95, 97	Nb	95
Mo	99	Ru	103, 106
Sb	127, 129	Te	129m, 131m, 132
I	131, 132, 133, 134, 135	Cs	134, 136, 137
Ba	140	La	140
Ce	141, 143, 144	Pr	143
Np	239	Pu	238
Cm	242		

Ingestion

I	131, 133	Cs	134, 136, 137

Table 2.3 Nuclides included in the DBA source term

Deposited γ dose

Mn	54	Co	58, 60
Zn	65	Zr	95
Ag	110m	Sb	124
I	131, 132, 133, 134, 135	Cs	134, 137
Ba	140		

Cloud γ dose

Mn	54	Co	58
Kr	88	Rb	88, 89
Sr	91	Ag	110m
Te	129, 131m, 131	I	131, 132, 133, 134, 135
Xe	133, 135, 138	Cs	134, 137, 138
Ba	139	La	140
W	187		

Skin

Co	59	Sr	89
Y	93	Mo	99
Te	129, 131m	I	131, 132, 133, 134, 135
Cs	134, 137	Ba	139, 140
La	140	W	187

Inhalation

Na	24	Mn	54
Fe	55, 59	Co	58, 60
Zn	65	Rb	86, 88, 89
Sr	89, 90, 91	Y	91, 93
Zr	95	Nb	95
Mo	99	Ru	103, 106
Ag	110m	Sb	124
Te	127m, 129, 129m, 131, 131m, 132		
I	131, 132, 133, 134, 135	Cs	134, 136, 137, 138
Ba	139, 140	La	140
Ce	141, 143, 144	W	187
Np	239	Pu	236, 238, 239, 240, 241
Am	241, 243	Cm	242, 244

Ingestion

Mn	54	Co	58, 60
Zn	65	Sr	90
Ag	110m	Te	129m
I	131, 133	Cs	134, 137
Ce	144		

3 EXPERT JUDGEMENT ASPECTS OF THE STUDY

The expert judgement aspects of the study were undertaken jointly by the USNRC and EC, using a combination of methods developed in earlier American[1] and European[2] studies. The formal use of expert judgement followed an agreed protocol, though this was modified during the elicitation process; the final version, which includes comments on possible improvements, is described in reference 3. This section summarises the information provided by the experts, the way in which the views of the different experts were combined, and the methods adopted to convert the distributions provided by the experts into distributions on the COSYMA input parameters.

Accident consequence programs consider the various pathways by which people can be irradiated following a release to atmosphere, together with the calculation of doses and risks following intakes of radionuclides. The input parameters are drawn from many scientific fields. Obtaining information on the uncertainty in the parameter values from expert judgement therefore also requires experts from the different scientific fields covered by the parameters, and no one group of people would have expert knowledge about the whole set of parameters. Therefore a series of expert panels was formed, with each panel covering a particular aspect of the overall area. The panels covered

- atmospheric dispersion,
- deposition,
- transfer through terrestrial food chains, split into two panels for soil and plant processes and for animal processes.
- external γ exposure from deposited material,
- internal dosimetry,
- risk of early health effects,
- risk of late health effects.

Each of the panels included about 8 experts, chosen according to an agreed set of criteria, and included both European and American experts. However, for the food chain study, separate panels of European and American experts were formed and some different questions were posed to the two panels as some conditions and agricultural practices are different in Europe and the US. The experts were introduced to the modelling used in accident consequence systems. They were also trained in assessing probability distributions. They were then asked for their views on the distributions of values for a number of parameters in the modelling area in which they are expert. In providing these distributions, the experts were free to use whatever models or information they wished. As mentioned in Section 2, the parameters for which distributions were elicited using expert judgement were considered to be those whose uncertainty is likely to make large contributions to the overall uncertainty, within the area of expertise of the panel. Distributions were obtained for the other parameters using less formal methods.

Some of the parameters in accident consequence models represent quantities that can, in principle, be measured. Others, for example some transfer coefficients used in food chain models, cannot and so must be derived from other measurable quantities. A fundamental aspect of the

methodology of formal expert judgement elicitation is that experts should only be asked to give their views on values of quantities that can potentially be measured. For example, the atmospheric dispersion experts were asked for information on the air concentration at particular distances, rather than on the parameters used in the dispersion models. In this way a library of information can be obtained that can be of use to both the MACCS and the COSYMA systems, and also to other models and programs. In some cases, experts were asked for information on quantities that are not used in models currently included in either MACCS or COSYMA if this information might be of interest in other studies.

The panels provided information on a range of quantities as summarised below.

The atmospheric dispersion panel[4] provided information on the air concentration at a series of points and the standard deviation of the cross-wind distribution of activity at selected distances, for a number of atmospheric conditions. The distributions for the parameters of the COSYMA dispersion model have been derived from this information.

The deposition panel[4] provided information directly on dry deposition velocities to different surfaces for iodine and for a range of particle sizes. It also gave information on the amount of material remaining in the atmosphere after periods of rain; this has been used to derive the distributions for the washout coefficient.

The food chain panels[5] gave information on the food chain model input parameters for strontium, caesium and iodine for transfer into crops and transfer into animals' meat and cows' milk. The distributions on the values for the parameters of the model used to generate the COSYMA food chain libraries were derived from this information.

The deposited material panel[6] gave information on the dose and dose rate at different times following unit deposition of particular radionuclides. The information was given for the dose over a large uniform area of grass and for an average urban area. The panel also gave information on the shielding properties of various types of buildings, and on some aspects of population behaviour.

The internal dosimetry panel[7] gave information on the amounts of material deposited in different organs at a series of times following intake by inhalation and ingestion, for a range of nuclides. Distributions of values for the parameters in the model to calculate dose coefficients used with COSYMA were derived from this information. They also gave distributions for the doses in selected organs for ^{90}Sr, ^{131}I, ^{132}Te, ^{137}Cs, ^{144}Ce and ^{239}Pu.

The early health effects panel[8] gave information on the D_{10}, D_{50} and D_{90} doses* for a series of dose rates, for both whole body exposure and for preferential exposure of particular organs. Distributions of values for the parameters of the dose response relationships adopted in COSYMA were derived from this information.

---

* D_{10}, D_{50}, and D_{90} are the doses which cause the early effect in 10, 50 and 90% of the exposed population, respectively

The late health effects panel[9] gave information on the numbers of cancers, and cancer deaths, in a population within different time periods after exposure at a high and a lower dose rate. Information was obtained for both an average population and for children.

Each member of the panel gave their views independently for each of the elicited quantities. They also gave information on the correlations between the values of the different parameters. The distributions from each expert were aggregated into a single distribution.

3.5 Combining distributions from different experts

The method adopted for eliciting expert judgement gave the probability distributions for the values of various quantities representing the views of each expert separately. The distributions had to be combined into a single distribution before the uncertainty analysis could be undertaken. The distributions from each expert were combined by simply summing them to give uncertainty distributions from the panel.

This procedure has been criticised by a reviewer of an early document describing one of the module analyses, who suggested that a better procedure would be to propagate each expert's views individually through the model, and then combine the distributions on the model output. This procedure was tried for a part of the food chain library, and the uncertainty distributions on concentrations in foods compared with those obtained by combining the distributions on model inputs. In this case, the resulting distributions were very similar. The input parameters whose uncertainties make large contributions to the overall uncertainty on concentrations in foods were also determined for both methods of combining the expert distributions. Again, the results were very similar for the two methods.

3.6 Eliciting and processing dependencies between variables

The values to be assigned to some of the parameters within their ranges are likely to be correlated. For example, the concentrations of different nuclides in milk at short times after the deposition may be correlated as the processes removing material from grass are likely to be equally applicable to different elements. It has long been known that significant errors in uncertainty analysis can be caused by ignoring dependencies between uncertainties[10]. New techniques for estimating and analyzing dependencies in uncertainty analysis have been developed in the course of the joint EC/USNRC accident consequence uncertainty analysis. This section describes how the various dependencies were elicited from experts and combined. More detail of the methods used is presented in Appendix C.

The experts within each panel were asked for information on the dependencies between the different quantities considered by that panel. Some of the quantities considered by different panels may also be correlated. In these cases, the correlations have been specified by the project staff or by individuals who were members of two expert panels. The justifications for the correlations chosen are described in the report on the overall uncertainty analysis.[11]

3.2.1 Eliciting dependencies

The best source of information about dependencies is often the experts themselves. The most thorough approach would be to elicit directly the experts' joint distributions. The practical drawbacks to this approach have forced analysts to look for other dependence elicitation strategies. One obvious strategy is to ask experts directly to assess a (rank) correlation coefficient. However, even trained statisticians have difficulty with this type of assessment task[12].

Within the joint EC/USNRC study a new strategy has been employed for eliciting dependencies from experts. When the analyst has identified a potential dependence between (continuous) variables X and Y, experts first assess their marginal distributions for X and Y. They are then asked:

Suppose Y were observed in a given case and its value were found to lie above the median value which you have suggested for Y; what is your probability that, in this case, X would also lie above the median value you have suggested for X?

An appropriate joint distribution on the values of X and Y was selected which has

- the assessed marginal distributions of X and Y,
- the probability specified by the expert for both quantities lying above their median.
- and has minimal information among all distributions satisfying the above.

In the soil\plant panel, the project staff identified a large number of potential dependencies. For two reasons it was decided that not all dependencies were to be elicited; firstly there would be too many questions and secondly, in eliciting all dependencies, it is almost impossible to assure that the resulting correlation matrix is positive definite (see appendix C for further details). Therefore it was decided that the experts were to be elicited on a selection of all possible dependencies. If this selection is such that the resulting dependency structure is an acyclic graph (see appendix C for further details), it is possible to find a joint distribution which

- satisfies the marginal distributions of the selected variables
- has a correlation matrix which is positive definite and satisfies the results as specified in the dependency structure
- and has minimal information among all distributions satisfying the above.

The dependency document was constructed by both the consequence analyst and the uncertainty analyst. The consequence analyst drafted a list with potential important dependencies which was then reviewed by the uncertainty analyst to see if the corresponding dependency structure would result in an acyclic dependency structure. After a number of iterations the final dependency document was given to the experts.

3.2.2 Combining dependencies

The method of combining dependencies is summarised here, and described in more detail in Appendix C.

In some cases, only one expert gave a value for the dependency between two variables. In this case, that value was taken to be the correlation between the distributions of the two quantities summed over the different experts.

A complication arises when information from different experts is combined. Each of the experts gave the probability that X and Y are above their median values as suggested by that expert, rather than above the median value of the combined distribution. Since the medians for *X* and *Y* were not the same for all experts, the conditional probabilities cannot simply be summed. Each expert's marginal distributions and correlation must be used to determine the probability that the expert would assign to both X and Y being above the median values of the combined distributions. These resulting probabilities can then be combined. The process is described in more detail in Appendix C and reference 13.

3.7 Obtaining distributions on the COSYMA input parameters from information provided by the experts

The experts were only asked to give distributions on the values of observable quantities. In some cases, the COSYMA input parameters are observable quantities, and so the experts can give information directly on these quantities. In other cases, the COSYMA input parameters are not observable and so the distributions on the input parameters must be obtained from distributions on related observable quantities. This section summarises the methods adopted for this process, which is termed "probabilistic inversion". Two computer programs, PARFUM and PREJUDICE, were developed for this process[14]. The PARFUM program was used in the early stages of the analysis, but was replaced by PREJUDICE in the later stages.

The quantities for which the experts provided information are termed "elicitation variables"; the COSYMA input parameters, for which distributions are to be derived, are termed "target variables". The process is illustrated here using an example from the atmospheric dispersion module. The variation of plume size σ with distance x is described using the COSYMA parameters a and b by the formula

$$\sigma = a\, x^b$$

Minimum and maximum values for consideration were specified by project staff for the parameters a and b. The values specified here need not be the values which are finally derived for the extremes of the distribution, but merely served to limit the ranges considered for the variables.

PARFUM generates a grid covering the region specified for the parameters in the target variable space. It then uses the model included in COSYMA to calculate the value for each of the

elicitation variables corresponding to each point of the grid in the target variables, as illustrated in Figure 3.1. The areas in the target variable space are then allocated probabilities so that the marginal distributions of all of the elicitation variables are reproduced as well as possible. PARFUM constructs the marginal distributions of the elicitation variables from the joint distribution of the target variables.

PREJUDICE uses a similar process, but considers the joint distribution of the elicitation variables. Again the model included in COSYMA is used to relate points in the target variable space with points in the elicitation variable space, as illustrated in Figure 3.2. The areas in the target variable space are then allocated probabilities so that the joint distribution of all of the elicitation variables is reproduced as well as possible. PREJUDICE constructs the joint distribution of the elicitation variables from the joint distribution of the target variables.

Further information on the general methods adopted is given in Appendix C. Information which is specific to the processing for each of the module analyses is presented in the reports on the particular module analysis.

3.8 References

1. US Nuclear Regulatory Commission. Severe accident risks; an assessment for five US nuclear power plants. Washington NUREG-1150 (1990)
2. Cooke R. Expert judgement study on atmospheric dispersion and deposition. Delft. TUD report 91-81 (1991)
3. Cooke R M, Goossens L H J, Kraan B C P. Probabilistic accident consequence uncertainty assessment - procedures guide using structured expert judgement. EUR 18820 (2000)
4. Harper F T, Hora S C, Young M L, Miller L A, Lui C H, McKay M D, Helton J C, Goossens L H J, Cooke R M, Päsler-Sauer J, Kraan B and Jones J A. Probabilistic accident consequence uncertainty analysis. Dispersion and deposition uncertainty assessment. NUREG/CR-6244, EUR 15855, SAND94-1453, Washington, DC/USA, and Brussels-Luxembourg, (1995)
5. Brown J, Goossens L H J, Kraan B C P, Cooke R M, Jones J A, Harper F T, Haskin F E, Abbott M L, Young M L, Hora S C, Rood A. Probabilistic accident consequence uncertainty analysis. Food chain uncertainty assessment. NUREG/CR-6523, EUR 16771, SAND97-0335 Washington, DC/USA, and Brussels-Luxembourg, (1997)
6. Goossens L H J, Boardman J, Kraan B C P, Cooke R M, Jones J A, Harper F T, Young M L and Hora S C. Probabilistic accident consequence uncertainty analysis: Uncertainty assessment for deposited material and external doses. Report NUREG/CR-6526, EUR 16772, Washington, DC/USA, and Brussels-Luxembourg, (1997)
7. Goossens L H J., Harrison J.D, Kraan B.C.P, Cooke R.M, Harper F.T. and Hora S.C. Probabilistic accident consequence uncertainty analysis: Uncertainty assessment for internal dosimetry, Report NUREG/CR-6571, EUR 16773, Washington, DC/USA, and Brussels-Luxembourg, (1997)
8. Haskin F.E., Harper F.T, Goossens L H J, Randall J, Kraan B.C.P and Grupa J.B. Probabilistic accident consequence uncertainty analysis: Early health effects uncertainty assessment,

Report NUREG/CR-6545, EUR 16775, Washington, DC/USA, and Brussels-Luxembourg, (1997)

9. Little M P, Muirhead C R, Goossens L H J, Kraan B C P, Cooke R M, Harper F T and Hora S C. Probabilistic accident consequence uncertainty analysis: Late health effects uncertainty assessment, Report NUREG/CR-6555, EUR 16774, Washington, DC/USA, and Brussels-Luxembourg, (1997)
10. Apostolakis G. and Kaplan, S. Pitfalls in risk calculations. Reliability Engineering 1981:2, 135-145.
11. Jones J A, Ehrhardt J, Goossens L H J, Brown J, Cooke R M, Fischer F, Hasemann I and Kraan B C P. Probabilistic accident consequence uncertainty assessment using COSYMA: Overall uncertainty analysis. EUR-18826, FZKA-6312 (2000)
12. Gokhale, D. and Press, S. Assessment of a prior distribution for the correlation coefficient in a bivariate normal distribution. J.R. Stat. Soc. A 1982: vol: 145 P.2, 237-249.
13. Cooke R M and Kraan B C P. Dealing with dependencies in uncertainty analysis. In P C Cacciabue and I A Papazoglou (eds). Probabilistic Safety Assessment and Management (ESREL 96 – PSAM III, June 24-28, 1996, Crete, Greece). Volume 1, Springer, 1996 pp 625 – 630
14. Krann B C P and Cooke R M. Post-processing techniques for the joint CEC-USNRC uncertainty analysis of accident consequence codes. Journal of Statistical Computation and Simulation 57 (1-4) 243-261, (1996)

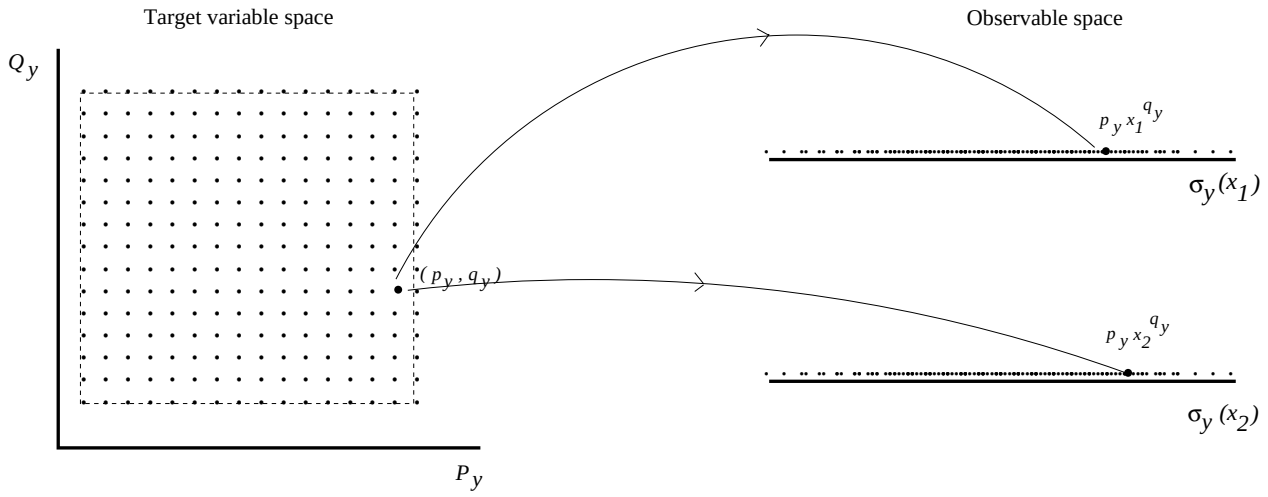


Figure 3.1 PARFUM mapping process to marginal distributions of the elicitation variables

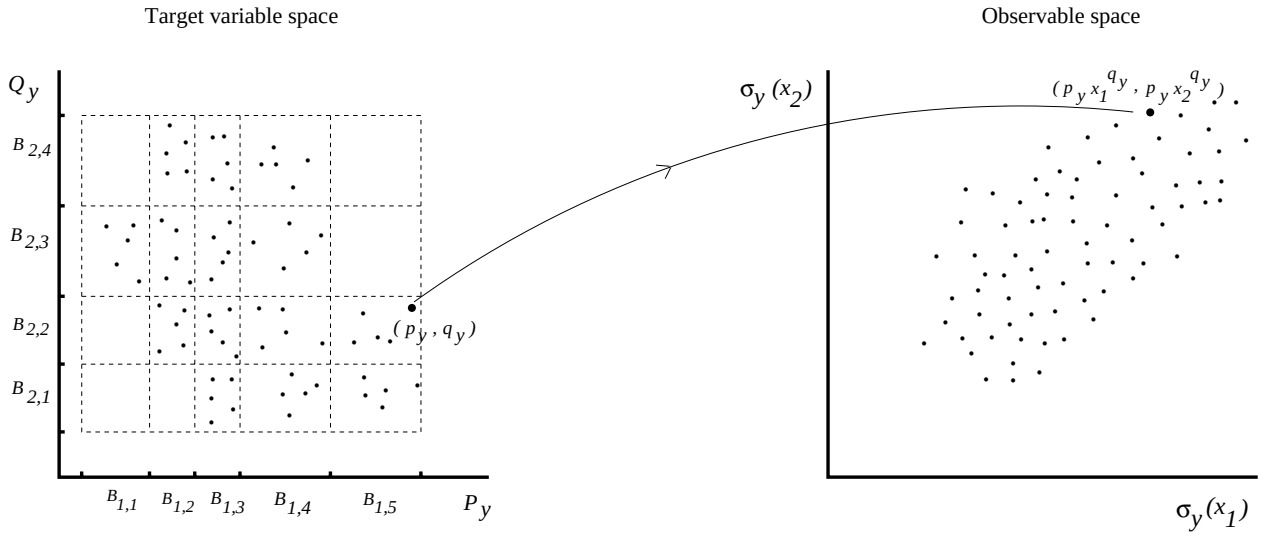


Figure 3.2 PREJUDICE mapping to joint distribution of the elicitation variables

4 METHOD OF UNDERTAKING THE ANALYSIS

The uncertainty in each of the modules was analysed and the parameters whose uncertainty make a major contribution to the overall uncertainty for that module were identified. This section describes the method used to undertake the analyses. It also describes the method adopted for determining which parameters should be included in the final analysis. Section 0 describes the method of sampling from the joint distribution over the input parameters. Section 0 describes the method of estimating the uncertainty on the COSYMA predictions. Section 0 describes the method of identifying the parameters whose uncertainties make major contributions to the overall uncertainty for each analysis, either of a particular module or of the overall system. Finally, section 0 describes the method adopted to determine which parameters from the module analyses should be included in the final analysis.

Sampling from the input distributions

The uncertainty on the input parameters has been expressed in terms of a joint distribution over the values of those parameters. Sets of values of the input parameters were obtained from this joint distribution by latin hypercube sampling (LHS) using the Sandia LHS program[(1)]. This program considers the joint distribution over the set of input parameters in terms of the marginal distribution of each of the input parameters and the rank correlation matrix between the values of the input parameters. The marginal distributions and rank correlations were extracted from the joint distribution.

Values were selected from the marginal distributions for each of the parameters using the LHS approach. LHS is a modified random sampling technique using a stratified sampling system, that reduces the variance of the uncertainty on the model output compared to other techniques for the same sample size. Simple random sampling could, by chance, choose all the parameter values from only a part of their complete range, and possibly omit values from part of the range that could greatly influence the predicted consequences.

The LHS procedure forces the selected values of each model parameter to be spread across its entire range. A latin hypercube sample of size *n* divides the range of each model parameter into *n* non-overlapping intervals of equal probability, and randomly selects a value from each of these intervals. If X_i is the set of values selected for the model parameters for one run of the program being analysed, then there will be n sets of values of X_i. The n values for X_1 are paired at random with the n values for X_2 which are then combined at random with the values for X_3. The process is continued until all the values X_i have been combined.

This process may produce unwanted correlations between model parameters within an LHS sample, due to the random pairings of the parameters as the sample is generated. Such correlations may introduce incorrect correlations between output model parameters or influence the extent of the predicted uncertainty. The Sandia LHS program includes a method, introduced by Iman and Conover[(2)], to reduce these correlations. The technique preserves the fundamental nature of LHS, but replaces the random pairing of model parameter values with a pairing that keeps all of

the pairwise rank correlations among the model parameters as close to zero as possible. The technique can also be used to induce a desired rank correlation structure among the model parameters. The procedure samples independently from the marginal distribution of each of the parameters, and then combines the sets of parameter values in the way that best reproduces the desired rank correlation structure between the input parameters.

Previous experience with uncertainty analyses of PRA codes suggests that the minimum sample size should be at least 1.5 times the number of uncertain input parameters[3]. Estimates of the uncertainty derived from different samples of this size are found to be reasonably similar. This sample size has also been found to give reasonable confidence in identifying the parameters whose uncertainties make major contributions to the overall uncertainty.

Estimation of extent of uncertainty

The next stage of the analysis is to run COSYMA for each of the sets of sampled input parameter values. The uncertainty is then obtained from the results obtained with the different sets of input parameters. The process is illustrated in Figure 4.1

COSYMA aims to evaluate the risks of future accidents. The atmospheric conditions at the time of the accident cannot be known in advance. However, the probability distribution of atmospheric conditions can be known, and a single run of COSYMA produces a probability distribution of each of the consequences considered, allowing for the uncertainty on the atmospheric conditions at the time of the accident. This distribution is described here as the ccdf (complementary cumulative distribution function) on the selected consequence. The ccdf can be described in terms of characteristic quantities, such as the expectation value and the value of various percentiles. Results are presented in this study for the mean value and the 95^{th} and 99^{th} percentiles of the distribution.

Because of the lack of knowledge of the most appropriate values to assign to the input parameters, there is uncertainty in the ccdf obtained. The uncertainty analysis, consisting of n runs of COSYMA, generated n ccdfs. The differences between the ccdfs from the different runs of COSYMA represents the uncertainty on the model predictions, which can be presented in different ways. The quantities used in this study to describe the uncertainty are summarised in Table **4.1**, and described below.

The result of any one run from the analysis is a ccdf reflecting the uncertainty because of the possible atmospheric conditions at the time of the accident. The ccdf generated using the default values of each of the input parameters is referred to in this study as the “reference curve”.

Since each of the input parameters has been assigned a value that may be correct, each of the output ccdfs also may be correct. The difference between the ccdfs from the different runs of the program represents the extent of the uncertainty reflecting the lack of knowledge on the correct value to assign to the different parameters.

The first method of presenting the uncertainty retains the different sets of ccdfs, which give n equally probable estimates of the expectation value and of each of the percentiles of the distribution. These values can be used to represent the uncertainty distribution on that endpoint of COSYMA. Various percentiles of the distribution on the output quantity (which may itself be a percentile of the ccdf for some endpoint) can then be extracted. In this way, the percentiles of the predicted uncertainty on the ccdf of consequences can be obtained. In the reports describing the module and final analyses, the uncertainty is characterised by the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles of the uncertainty distribution on each of the endpoints considered. The general discussion of the extent of the uncertainty is presented using the ratio of the 95th to 5th percentiles; this quantity is termed the "uncertainty factor". There is also an interest in the extent to which predictions obtained using the default value for each input parameter could underestimate the results. Therefore the ratio of the 95th percentile of the uncertainty distribution to the value obtained with the default values for the input parameters was also obtained. This quantity is termed the "reference uncertainty coefficient". The percentiles of the predicted uncertainty on the probability distribution can be joined to form an envelope that defines selected confidence bounds on the ccdf of predicted consequences. These envelopes are not, in general, the ccdf produced from any one run of COSYMA, corresponding to one set of input parameter values. In this study the 5%-envelope and the 95%-envelope are used in presenting the results. The region between these envelopes is the area with 90% confidence of containing the true reference curve which expresses the probability of exceeding particular levels of consequences allowing for the range of atmospheric conditions that could occur at the time of the accident.

This method of presenting the uncertainty separates the contribution from the uncertainty in the weather conditions at the time of the accident and the contribution from the uncertainty to assign to the different parameter values. It shows the extent to which the uncertainty could be reduced by further research into the values of the different parameters.

The second method combines the different ccdfs for each endpoint into a single probability distribution including the variations caused by the weather and the variations caused by the parameter uncertainties. This one overall ccdf gives the uncertainty distribution of endpoint values caused by all quantifiable uncertainties: those coming from the range of atmospheric conditions at the time of the accident and those resulting from parameter uncertainties. This form of presentation of the results does not give any information on the extent to which the uncertainty could be reduced by further research. In this study, this ccdf is referred to as the "mean curve". Some of the results from the overall analysis are presented using the ratio of the 95% envelope to the mean curve. This quantity is termed the "mean uncertainty coefficient".

The results from the module analyses are mainly presented using the first method described here. The results from the overall analysis are presented using both methods described here.

Identifying the important parameters from a single analysis

The final stage of a single analysis, either a module analysis or the final analysis of

COSYMA, is to identify those input parameters whose uncertainties make major contributions to the overall uncertainty in the predicted consequences. This can only be undertaken while the results from the different runs of COSYMA are kept separate (i.e. before the results are combined to give the 5% and 95% envelopes or the mean curve). The method adopted in this analysis to identify the important uncertainties uses a combination of partial rank correlation and linear regression.

The partial correlation coefficient (PCC) is a measure that explains the linear relationship between an input parameter value and the predicted consequence when the possible linear effects of the remaining parameters have been removed. In this analysis, the partial rank correlation coefficient (PRCC) was calculated; this considers correlations between the ranks of the input and output quantities, rather than the actual values, and so examines the strength of monotonic relationships between the input and output quantities when the effects of monotonic relationships between the output and any of the other inputs have been removed. The magnitude of the PRCC reflects the strength of the correlation, while its sign shows whether the model output increases or decreases as the value of the input parameter increases.

Random effects may mean that there is an apparent correlation between an input and output parameter which are in fact not correlated. Fischer[4] has proposed a method of determining the value of PRCC that could be obtained by chance for a particular sample size and number of runs. He shows that the "null" hypothesis of no correlation should be rejected if the magnitude of

$$T = \frac{t_{\alpha/2,n-k}}{\sqrt{n-k+t^2_{\alpha/2,n-k}}}$$

the PRCC is greater than T,

where n is the number of COSYMA runs,

k is the number of uncertain input parameters,

$t_{\alpha/2,n-k}$ is the $(1-\alpha/2)$-quantile of the t-distribution with n-k degrees of freedom,

and α is the probability that the particular value of T could occur by chance.

Those input parameters for which the absolute value of the PRCC is greater than T are assumed to make a contribution to the uncertainty on the particular endpoint considered. The input parameters are ordered in terms of the absolute values of the PRCC to identify the most important parameters for a particular (module or overall) analysis.

The contribution of each parameter to the uncertainty on an endpoint can be determined using the "coefficient of determination", normally denoted R^2. This quantity reflects the fraction of the uncertainty that can be explained by linear relationships between the input parameters and the model output. The percentage contribution of a particular parameter to the overall uncertainty, P, can be expressed as

$$P = 100\ R^2_p/R^2_t$$

where: R^2_p is the coefficient of determination obtained when only the particular parameter is

included in the regression analysis

and: R^2_t is the coefficient of determination obtained when all input parameters are included in the regression analysis.

Regression analysis combines the effects of correlated variables into the value of R^2, and hence P, so that the apparent contribution of a parameter to the overall uncertainty calculated in this way can include some of the contributions from input parameters whose value is correlated to that of the parameter being considered. For this reason, the percentage contributions of the uncertainty on the input parameters to the overall uncertainty may total a value that is greater than or less than 100%, depending on whether the correlations are positive or negative. It also means that parameters with large values of PRCC could appear to make only small percentage contributions to the uncertainty, or that parameters with small values of PRCC could appear to make large percentage contributions to the uncertainty. Consider the following examples to illustrate this problem.

First, assume that the model is represented by

$$y = x_1^2 + k\, x_2\, (x_3 + x_4)$$

where: y is the model output, x_i are parameters and k is a constant.

Suppose that the values of the parameters and constant are such that the first term is much larger than the second term. Now consider the case when the uncertainty on x_1 is large, that on x_4 is small and that on x_2 and x_3 are similar but smaller than that on x_1, and that the parameters are uncorrelated. It is likely that an analysis using PRCC's will identify x_1 as the only parameter making a major contribution to the overall uncertainty; x_4 may be found to make a negligible contribution to the uncertainty. Now consider the case where there is a substantial correlation between x_1 and x_4. PRCC's do not consider the effects of the correlations, and so should again identify x_4 as making a negligible contribution. However, the correlations could affect the calculation of percentage contributions to the uncertainty so that x_4 is identified as important in a regression analysis. However, in this situation it seems reasonable to assume that the apparent importance of x_4 reflects the actual importance of x_1 and the correlations between these two parameters. Therefore x_4 would not be considered to be an important parameter in this situation.

Now consider the simple case where the model can be represented by

$$y = x_1 + x_2 + x_3$$

Suppose that the uncertainties on x_1 and x_2 are similar and much larger than that on x_3, and that the parameters are not correlated. It is likely that an analysis using either PRCCs or percentage contributions will identify both x_1 and x_2 as making major contributions to the overall uncertainty. Now consider the case where there is a large and negative correlation between these two parameters. Again, an analysis using PRCCs should identify both parameters as making a large contribution to the overall uncertainty as they do remove the effects of correlations. However, the

correlations could again affect the calculation of percentage contribution; it is possible that neither parameter would appear as important in this case.

Selection of parameters for the overall analysis

The analysis of the uncertainty in COSYMA has been undertaken in stages. The first stage was to analyse the uncertainty in the predictions of COSYMA reflecting uncertainty in the input parameters for a particular part of the complete model - termed “module analyses”. These analyses also identified the parameters whose uncertainty make a large contribution to the overall uncertainty from that module. The final stage was to combine the parameters that are identified as important from the module analyses into a single, overall analysis of the uncertainty from the complete system. There is no accepted method of determining which uncertain parameters should be included in the overall analysis. This section describes the problems involved in the choice and the procedure selected.

The analysis methods used in this study provide information on the partial rank correlation coefficients (PRCC) between the various input and output quantities, and also provide information on the fraction of the total uncertainty generated by the uncertainty on each parameter separately. This gives contributions to the uncertainty from each of the modules separately, but does not provide a simple method of comparing the contributions of parameters in different modules.

A series of analyses was made of the extent of the uncertainty when different numbers of parameters were included in the analysis. This was intended to identify a method of selecting a reduced number of uncertain parameters that give a reasonable representation of the overall uncertainty with all the parameters considered in the module analysis. These analyses were based on the dispersion and deposition and the health effects module analyses, as these do not require the generation of large numbers of data libraries.

Some of the results from the dispersion and deposition investigations are presented in Table 4.2 to Table 4.5. The tables present the results for the air concentration and deposition of elemental iodine and of caesium at three distances, for different methods of reducing the number of uncertain parameters considered. The selection procedures used were as follows:

- For Table 4.2, the parameters included were those with first or second rank according to PRCC for at least one endpoint. Other parameters were given their default value.
- For Table 4.3, the parameters included were those with first rank according to PRCC for at least one endpoint plus any parameter for which the correlation coefficient with any of the selected parameters is greater than 0.3. Other parameters were set at their default value.
- For Table 4.4, the parameters included were those with first rank according to PRCC for at least one endpoint plus any parameter for which the correlation coefficient with any of the selected parameters is greater than 0.3. Other parameters were set at the median value of their distribution from the analysis.
- For Table 4.5, the parameters included were those that made a percentage contribution of more than 15%. Other parameters were set at the median value of their distribution from the

analysis.

The parameters included in the analyses shown in these tables were those making the appropriate contribution to the uncertainty for at least one quantity considered (ie mean value, 95th or 99th percentile) for at least one endpoint considered.

The 5^{th} and 95^{th} percentiles of the uncertainty distribution, and the uncertainty factor, were calculated for the different endpoints using the full set of uncertain parameters and using the reduced set of parameters. The tables present the results only for the uncertainty on the mean value of the endpoints.

The ratio of the 5th percentile of the uncertainty distribution, calculated using the full set of parameters, to that calculated using the reduced set of parameters was obtained for each endpoint, and the average value of the ratio calculated. This process was also repeated for the 95th percentile of the uncertainty distribution, and for the "uncertainty factor". The average values of these ratios are also presented in the tables. These results show that, in this situation, it is possible to identify a sub-set of input parameters that give a reasonable representation of the overall uncertainty, and that all subsets considered give a reasonable representation of the 95^{th} percentile of the distribution on the endpoints, with a less accurate representation of the 5^{th} percentile of the uncertainty distribution.

The possibility of reducing the number of parameters was also considered using the health effects module results. Here, all of the uncertain parameters contributing to early health effects are identified as important for at least one of the endpoints considered, and so would be selected for the final analysis. A different situation was found for the uncertainty on the results for late health effects. The numbers of cancers in each organ are similar, and the total number of cancers is the sum of those numbers. Omitting the risk coefficient for any organ means that the uncertainty on the number of cancers in that organ is not considered in the total uncertainty. In this situation, it is not possible to reduce the number of parameters considered in the analysis. There is no other part of COSYMA where the predicted consequence is the sum of several quantities that have a similar value, and so this situation does not invalidate the argument that a sub-set of parameters can be identified for the overall analysis from the other module analyses.

These results show that a limited number of parameters can be chosen that will give a good representation of the upper and lower ends of the uncertainty distribution, other than in cases where the model amounts to a sum of terms of similar magnitude. While the selection of those parameters in the first and second ranks gave the best results here, there is no reason to suppose that this will always give the best results in all situations.

The option of including the parameters ranked in first and second place according to their PRCC together with those making more than 15% contribution to the uncertainty was considered. It is reasonable to assume that, if the overall analysis includes the parameters whose uncertainties are important for each module separately, then it will automatically include the parameters whose uncertainties are important for the complete system.

The module analyses have shown that analyses with about 160 uncertain parameters are feasible. The dose and food chain modules have larger numbers of input parameters than the dispersion and health effects modules, and so it is more important to reduce the number of parameters selected from those two modules. Therefore the possibility of selecting a similar number of parameters for the overall analysis was investigated, based on a combination of the selection procedures adopted for the results presented in Table 4.2 to Table 4.5 for the food chain and the dose module parameters. Those parameters in the first two ranks according to PRCCs and those parameters whose uncertainty makes provides than 15% of the overall uncertainty were considered. This process identified about 40 parameters from the food chain module and about 60 parameters from the dose module. The numbers were increased slightly by including all parameters from any compartment models for which some of the parameters were identified. The dispersion and deposition module analysis included less than 30 parameters, of which four relate to the deposition of organic iodine. The fraction of iodine released in the organic form is very small (about 1%) and so these parameters can be omitted as this form of iodine makes only a very small contribution to the predicted consequences and the uncertainty on them. The health effects module also considered about 30 parameters. As explained above, it is not possible to identify a sub-set that gives a reasonable representation of the overall uncertainty for this module, so all the health effects parameters were selected. This provided a selection of about 180 parameters for inclusion in the overall analysis.

References

1. Iman R L and Shortencarier M J. A fortran 77 program and user's guide for the generation of latin hypercube and random samples for use with computer models. NUREG/CR-3624 (1984)

2. Iman R L and Conover W J. A distribution free approach to inducing rank correlation among input variables. Comm. Statist. Comput, 11 311-334 (1982)

3. Fischer F. Procedures for uncertainty analyses of UFOMOD. A user guide. KfK 4626 (1990)

4. Fischer F, Ehrhardt J and Hasemann I. Uncertainty and sensitivity analyses of the complete program system UFOMOD and of selected submodels. Karlsruhe, KfK 4627 (1990)

Table 4.1 Quantities used to present the extent of the uncertainty

Results presented as curves or probability distributions

Term used in the reports	Definition
Reference curve	ccdf calculated using default values for all parameters, presents the probability that consequences will be greater than a particular value allowing for the range of possible atmospheric conditions at the time of the accident
95% envelope	ccdf with 95% confidence that the true reference curve will be lower than this value
5% envelope	ccdf with 95% confidence that the true reference curve will be higher than this value
Mean curve	ccdf the expresses the probability that consequences will be greater than a particular value allowing for both the uncertainty on the most appropriate value to assign to the input parameters and the range of atmospheric conditions at the time of the accident

Results presented as ratios

Term used in the reports	Definition
Uncertainty factor	Ratio of the 95th to5th percentiles of the uncertainty distribution
Reference uncertainty coefficient	Ratio of the 95th percentile of the uncertainty distribution to the reference value
Mean uncertainty coefficient	Ratio of the 95th percentile of the uncertainty distribution to the value obtained from the mean curve

Table 4.2 Comparison of uncertainties using all uncertain parameters and those in first or second rank according to PRCC, other parameters given default values

Endpoint considered	Results with reduced number of uncertain parameters			Results with all uncertain parameters		
	5th percentile	95th percentile	Uncertainty factor	5th percentile	95th percentile	Uncertainty factor
Deposition of iodine at 0.875 km	4.23E+09	1.57E+11	3.71E+01	4.64E+09	1.49E+11	3.20E+01
Deposition of iodine at 4.9 km	5.73E+08	4.55E+09	7.94E+00	5.68E+08	3.75E+09	6.60E+00
Deposition of iodine at 20 km	3.94E+07	1.95E+08	4.96E+00	4.12E+07	1.97E+08	4.79E+00
Deposition of caesium at 0.875 km	1.95E+07	3.72E+09	1.91E+02	2.14E+07	3.26E+09	1.52E+02
Deposition of caesium at 4.9 km	2.81E+06	1.87E+08	6.67E+01	3.99E+06	1.52E+08	3.80E+01
Deposition of caesium at 20 km	5.00E+05	9.63E+06	1.93E+01	4.60E+05	1.04E+07	2.26E+01
Air concentration of iodine at 0.875 km	9.77E+11	8.70E+12	8.91E+00	9.27E+11	7.89E+12	8.51E+00
Air concentration of iodine at 4.9 km	2.34E+10	9.31E+11	3.97E+01	2.06E+10	7.92E+11	3.84E+01
Air concentration of iodine at 20 km	1.22E+09	6.25E+10	5.11E+01	7.56E+08	6.86E+10	9.07E+01
Air concentration of caesium at 0.875 km	9.03E+10	5.42E+11	6.00E+00	8.29E+10	4.34E+11	5.23E+00
Air concentration of caesium at 4.9 km	7.96E+09	7.24E+10	9.09E+00	7.04E+09	6.00E+10	8.52E+00
Air concentration of caesium at 20 km	4.23E+08	5.86E+09	1.39E+01	4.51E+08	6.33E+09	1.40E+01

Average value of ratio of 5th percentile for all parameters to that for reduced parameter list is 0.99
Average value of ratio of 95th percentile for all parameters to that for reduced parameter list is 0.93
Average value of ratio of uncertainty factors for all parameters to that for reduced parameter list is 0.98

Table 4.3 Comparison of uncertainties using all uncertain parameters and those in first rank according to PRCC together with correlated parameters, other parameters given default values

Endpoint	Results with reduced number of uncertain parameters			Results with all uncertain parameters		
	5th percentile	95th percentile	Uncertainty factor	5th percentile	95th percentile	Uncertainty factor
Deposition of iodine at 0.875 km	1.25E+09	1.52E+11	1.22E+02	4.64E+09	1.49E+11	3.20E+01
Deposition of iodine at 4.9 km	2.27E+08	3.92E+09	1.73E+01	5.68E+08	3.75E+09	6.60E+00
Deposition of iodine at 20 km	3.92E+07	2.08E+08	5.30E+00	4.12E+07	1.97E+08	4.79E+00
Deposition of caesium at 0.875 km	2.21E+07	3.25E+09	1.47E+02	2.14E+07	3.26E+09	1.52E+02
Deposition of caesium at 4.9 km	3.27E+06	1.67E+08	5.11E+01	3.99E+06	1.52E+08	3.80E+01
Deposition of caesium at 20 km	4.47E+05	1.12E+07	2.50E+01	4.60E+05	1.04E+07	2.26E+01
Air concentration of iodine at 0.875 km	1.16E+12	8.08E+12	6.99E+00	9.27E+11	7.89E+12	8.51E+00
Air concentration of iodine at 4.9 km	2.18E+10	9.05E+11	4.15E+01	2.06E+10	7.92E+11	3.84E+01
Air concentration of iodine at 20 km	1.04E+09	7.40E+10	7.14E+01	7.56E+08	6.86E+10	9.07E+01
Air concentration of caesium at 0.875 km	1.01E+11	4.29E+11	4.25E+00	8.29E+10	4.34E+11	5.23E+00
Air concentration of caesium at 4.9 km	8.30E+09	6.09E+10	7.34E+00	7.04E+09	6.00E+10	8.52E+00
Air concentration of caesium at 20 km	4.57E+08	5.64E+09	1.23E+01	4.51E+08	6.33E+09	1.40E+01

Average value of ratio of 5th percentile for all parameters to that for reduced parameter list is 1.3
Average value of ratio of 95th percentile for all parameters to that for reduced parameter list is 0.97
Average value of ratio of uncertainty factors for all parameters to that for reduced parameter list is 0.93

Table 4.4 Comparison of uncertainties using all uncertain parameters and those in first rank according to PRCC together with correlated parameters, other parameters given median values

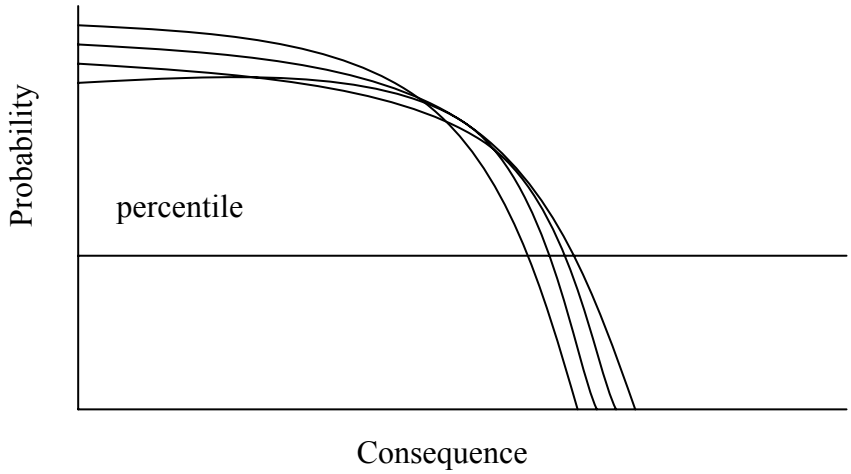
Endpoint	Results with reduced number of uncertain parameters			Results with all uncertain parameters		
	5th percentile	95th percentile	Uncertainty factor	5th percentile	95th percentile	Uncertainty factor
Deposition of iodine at 0.875 km	9.63E+08	1.33E+11	1.38E+02	4.64E+09	1.49E+11	3.20E+01
Deposition of iodine at 4.9 km	1.80E+08	3.32E+09	1.84E+01	5.68E+08	3.75E+09	6.60E+00
Deposition of iodine at 20 km	3.42E+07	1.80E+08	5.26E+00	4.12E+07	1.97E+08	4.79E+00
Deposition of caesium at 0.875 km	2.08E+07	3.27E+09	1.57E+02	2.14E+07	3.26E+09	1.52E+02
Deposition of caesium at 4.9 km	3.26E+06	1.46E+08	4.49E+01	3.99E+06	1.52E+08	3.80E+01
Deposition of caesium at 20 km	4.75E+05	1.02E+07	2.14E+01	4.60E+05	1.04E+07	2.26E+01
Air concentration of iodine at 0.875 km	7.89E+11	7.75E+12	9.83E+00	9.27E+11	7.89E+12	8.51E+00
Air concentration of iodine at 4.9 km	2.13E+10	7.67E+11	3.60E+01	2.06E+10	7.92E+11	3.84E+01
Air concentration of iodine at 20 km	1.18E+09	6.33E+10	5.36E+01	7.56E+08	6.86E+10	9.07E+01
Air concentration of caesium at 0.875 km	7.30E+10	4.10E+11	5.62E+00	8.29E+10	4.34E+11	5.23E+00
Air concentration of caesium at 4.9 km	6.49E+09	5.57E+10	8.59E+00	7.04E+09	6.00E+10	8.52E+00
Air concentration of caesium at 20 km	4.48E+08	5.08E+09	1.13E+01	4.51E+08	6.33E+09	1.40E+01

Average value of ratio of 5th percentile for all parameters to that for reduced parameter list is 1.53
Average value of ratio of 95th percentile for all parameters to that for reduced parameter list is 1.08
Average value of ratio of uncertainty factors for all parameters to that for reduced parameter list is 0.93

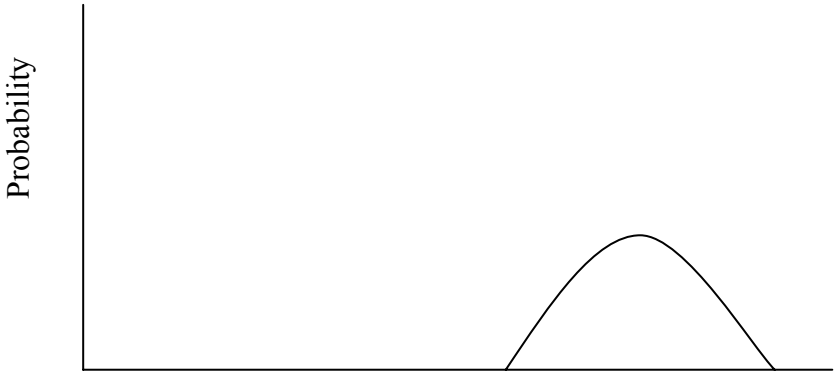
Table 4.5 Comparison of uncertainties using all uncertain parameters and those making more than 15% contribution to the uncertainty, with other parameters given median values.

Endpoint	Results with reduced number of uncertain parameters			Results with all uncertain parameters		
	5th percentile	95th percentile	Uncertainty factor	5th percentile	95th percentile	Uncertainty factor
Deposition of iodine at 0.875 km	9.58E+08	1.38E+11	1.44E+02	4.64E+09	1.49E+11	3.20E+01
Deposition of iodine at 4.9 km	1.80E+08	3.40E+09	1.89E+01	5.68E+08	3.75E+09	6.60E+00
Deposition of iodine at 20 km	3.25E+07	1.84E+08	5.67E+00	4.12E+07	1.97E+08	4.79E+00
Deposition of caesium at 0.875 km	2.19E+07	3.23E+09	1.48E+02	2.14E+07	3.26E+09	1.52E+02
Deposition of caesium at 4.9 km	3.24E+06	1.49E+08	4.59E+01	3.99E+06	1.52E+08	3.80E+01
Deposition of caesium at 20 km	4.67E+05	1.02E+07	2.19E+01	4.60E+05	1.04E+07	2.26E+01
Air concentration of iodine at 0.875 km	8.08E+11	7.96E+12	9.86E+00	9.27E+11	7.89E+12	8.51E+00
Air concentration of iodine at 4.9 km	2.27E+10	7.95E+11	3.49E+01	2.06E+10	7.92E+11	3.84E+01
Air concentration of iodine at 20 km	1.17E+09	6.65E+10	5.68E+01	7.56E+08	6.86E+10	9.07E+01
Air concentration of caesium at 0.875 km	8.00E+10	4.10E+11	5.13E+00	8.29E+10	4.34E+11	5.23E+00
Air concentration of caesium at 4.9 km	6.82E+09	5.73E+10	8.41E+00	7.04E+09	6.00E+10	8.52E+00
Air concentration of caesium at 20 km	4.48E+08	5.60E+09	1.25E+01	4.51E+08	6.33E+09	1.40E+01

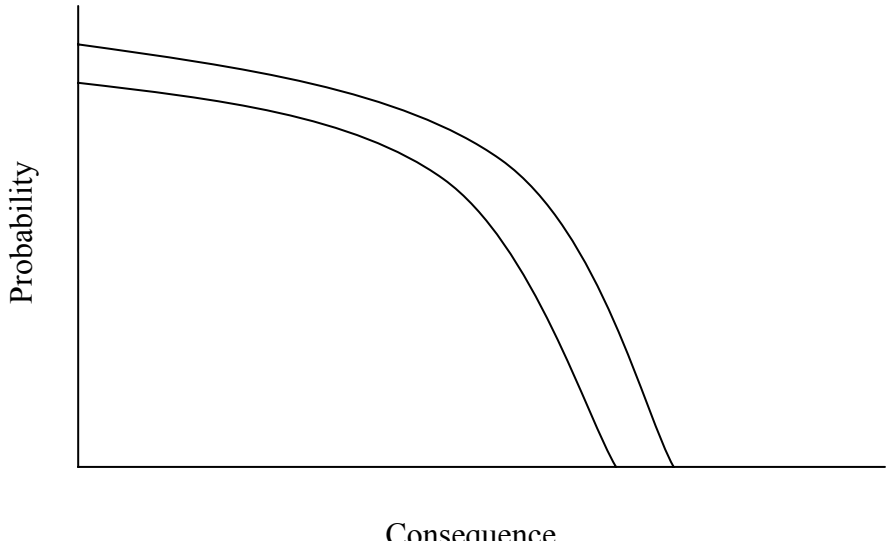
Average value of ratio of 5th percentile for all parameters to that for reduced parameter list is 1.52
Average value of ratio of 95th percentile for all parameters to that for reduced parameter list is 1.05
Average value of ratio of uncertainty factors for all parameters to that for reduced parameter list is 0.92



Step 1: derive many sets of ccdf's from the compiled sets of input



Step 2: determine the pdf on a percentile of the ccdf



Step 3: derive the envelope of the ccdf at a chosen confidence interval

Figure 4.1 Steps in determining uncertainty on model predictions

APPENDIX A

Reports from the Project

Reports on the expert elicitation

Harper F T, Hora S C, Young M L, Miller L A, Lui C H, McKay M D, Helton J C, Goossens L H J, Cooke R M, Päsler-Sauer J, Kraan B and Jones J A. Probabilistic accident consequence uncertainty analysis. Dispersion and deposition uncertainty assessment. NUREG/CR-6244, EUR 15855, SAND94-1453, Washington, DC/USA, and Brussels-Luxembourg, (1995).

Brown J, Goossens L H J, Kraan B C P, Cooke R M, Jones J A, Harper F T, Haskin F E, Abbott M L, Young M L, Hora S C, Rood A. Probabilistic accident consequence uncertainty analysis. Food chain uncertainty assessment. NUREG/CR-6523, EUR 16771, SAND97-0335 Washington, DC/USA, and Brussels-Luxembourg, (1997).

Goossens L H J, Boardman J, Kraan B C P, Cooke R M, Jones J A, Harper F T, Young M L and Hora S C. Probabilistic accident consequence uncertainty analysis: Uncertainty assessment for deposited material and external doses. Report NUREG/CR-6526, EUR 16772, Washington, DC/USA, and Brussels-Luxembourg, (1997).

Goossens L H J., Harrison J.D, Kraan B.C.P, Cooke R.M, Harper F.T. and Hora S.C. Probabilistic accident consequence uncertainty analysis: Uncertainty assessment for internal dosimetry, Report NUREG/CR-6571, EUR 16773, Washington, DC/USA, and Brussels-Luxembourg, (1997).

Little M P, Muirhead C R, Goossens L H J, Kraan B C P, Cooke R M, Harper F T and Hora S C. Probabilistic accident consequence uncertainty analysis: Late health effects uncertainty assessment, Report NUREG/CR-6555, EUR 16774, Washington, DC/USA, and Brussels-Luxembourg, (1997).

Haskin F.E., Harper F.T, Goossens L H J, Randall J, Kraan B.C.P and Grupa J.B. Probabilistic accident consequence uncertainty analysis: Early health effects uncertainty assessment, Report NUREG/CR-6545, EUR 16775, Washington, DC/USA, and Brussels-Luxembourg, (1997).

Goossens L H J, Jones J A, Ehrhardt J, Kraan B C P. Probabilistic accident consequence uncertainty assessment: Countermeasures Uncertainty Assessments. EUR 18821 and FZKA 6307 (2000).

Reports on the COSYMA uncertainty analysis

Cooke R M, Goossens L H J, Kraan B C P. Probabilistic accident consequence uncertainty assessment - procedures guide using structured expert judgement. EUR 18820 (2000).

Jones J A, Ehrhardt J, Fischer F, Hasemann I, Goossens L H J, Kraan B C P, Cooke R M. Probabilistic accident consequence uncertainty assessment using COSYMA: Uncertainty from the Atmospheric Dispersion and Deposition Module. EUR 18822 and FZKA 6308 (2000).

Brown J, Jones J A, Fischer F, Hasemann I, Goossens L H J, Kraan B C P, Cooke R M. Probabilistic accident consequence uncertainty assessment using COSYMA: Uncertainty from the Food Chain Module. EUR 18823 and FZKA 6309 (2000).

Jones J A, Fischer F, Hasemann I, Goossens L H J, Kraan B C P, Cooke R M. Probabilistic accident consequence uncertainty assessment using COSYMA: Uncertainty from the Health Effects Module. EUR 18824 and FZKA 6310 (2000).

Jones J A, Fischer F, Hasemann I, Goossens L H J, Kraan B C P, Cooke R M. Phipps A, and Khursheed A. Probabilistic accident consequence uncertainty assessment using COSYMA: Uncertainty from the dose module. EUR18825 and FZKA 6311 (2000).

Jones J A, Ehrhardt J, Goossens L H J, Fischer F, Hasemann I, Kraan B C P, Cooke R M. Probabilistic accident consequence uncertainty assessment using COSYMA: Overall uncertainty analysis. EUR 18826 and FZKA 6312 (2000).

Jones J A, Kraan B C P, Cooke R M, Goossens L H J, Fischer F, and Hasemann I. Probabilistic accident consequence uncertainty assessment using COSYMA: Methodology and processing techniques. EUR 18827 and FZKA 6313 (2000).

APPENDIX B

Summary of the COSYMA Accident Consequence Code

COSYMA is intended for probabilistic calculations of the off-site consequences of hypothetical accidental releases of radioactive material to atmosphere at nuclear sites. It calculates the health effects, impact of countermeasures and economic costs of the releases. The processes considered in the calculations, and the routes of exposure following accidental releases to atmosphere, are illustrated in Figure B.1. The calculation is divided into a number of steps, as is also illustrated in Figure 1. COSYMA is a modular code, with different modules addressing the different stages of the calculation. However, while Figure 1 illustrates the steps in the calculation, the modules of the codes do not correspond exactly with the boxes shown in that figure. The following sections give brief descriptions of the models included in COSYMA. In some cases, COSYMA includes more than one model for a particular feature. This appendix also specifies which of the models was used for this uncertainty analysis.

COSYMA was developed by the National Radiological Protection Board (NRPB) of the UK and Forschungszentrum Karlsruhe (FZK) of Germany, as part of the European Commission's MARIA project[1]. It represents a fusion of ideas from the NRPB program MARC[2], the FZK program system UFOMOD[3] and input from other MARIA contractors. The program package was first made available in 1990 for use on mainframe computers, and several updates have been released since then. A PC version was first released in 1993 and has since been updated[4].*

COSYMA is a package of programs and data bases, rather than a single program. The mainframe version contains three main accident consequence assessment programs together with a number of preprocessing and evaluation programs. The three main sub-systems of COSYMA are known as the NE, NL and FL sub-systems. The NE (near, early) sub-system is limited to calculating early health effects and the influence of emergency actions to reduce those effects and is intended for use in the region near to the site. The NL (near, late) subsystem is limited to calculating late health effects and the associated countermeasures, and is intended mainly for use in the region near to the site. The FL (far, late) sub-system is concerned with calculating late health effects and appropriate countermeasures at larger distances from the site. Each of these programs is further sub-divided into a series of modules for the various steps in the calculation. PC COSYMA incorporates the NE and NL sub-systems of the mainframe version.

The main endpoints of COSYMA are the numbers of health effects, the impact of countermeasures and the economic costs resulting from an accidental release. A large number of intermediate results are obtained in the process of calculating the major endpoints; these results include activity concentrations, individual and collective doses and the countermeasures that would be imposed

* The mainframe and PC versions of COSYMA are made available on behalf of the European Commission. People wishing to obtain the mainframe version of the system should contact Dr J Ehrhardt, FZK, Germany (e-mail RODOS@RODOS.FZK.DE; those wishing to obtain the PC version of the system should contact Dr J A Jones, NRPB, UK (e-mail Arthur.Jones@NRPB.ORG.UK).

at different locations. The package contains a series of evaluation programs that allow these results to be presented in a variety of ways.

Following an accidental release to atmosphere, people can be irradiated by a number of routes of exposure. The ones considered in COSYMA are:-

- external γ irradiation from material in the plume,
- external γ irradiation from material deposited on the ground
- external β irradiation of skin from material deposited on skin and clothes
- internal irradiation following the inhalation of material from the plume or of material that has been deposited and subsequently resuspended
- internal irradiation from the ingestion of contaminated foods.

COSYMA includes some models directly within the various modules or subsidiary programs, but in other cases it uses results of models taken from data libraries. Thus the atmospheric dispersion models are used directly. COSYMA does not however, include models for the contamination of food or dosimetric calculations, using instead data libraries giving the results of other models, which are not part of COSYMA, itself, but whose uncertainty is considered within the current study.

B.1 Atmospheric dispersion and deposition

Mainframe COSYMA contains five different models of atmospheric dispersion that are appropriate for different applications or are based on different assumptions and approximations[5].

The NE and NL sub-system include the MUSEMET[6] model, which was originally written at Forschungsanlage Julich but has been extensively modified at FZK for use with COSYMA. This is a segmented Gaussian plume model allowing for changes of atmospheric conditions and wind direction during plume travel. This model derives the sequences of atmospheric conditions affecting the plume from a data file giving hourly averages for wind speed and direction, stability category, precipitation intensity and mixing layer depth. It allows for the effects on the subsequent dispersion of plume rise and buildings near the release point. It also includes the effects of wet and dry deposition of the dispersing material. This model is also included in PC COSYMA.

The NE and NL sub-systems can also be used with the COSGAP or RIMPUFF dispersion models, which are provided as separate programs. COSGAP[7] is a Gaussian plume dispersion model, which is similar to MUSEMET but does not consider changes of wind direction during plume travel. It is based on the dispersion model in MARC. RIMPUFF[8], developed by Risø National Laboratory, Denmark, is a Gaussian puff trajectory model which derives the atmospheric conditions affecting the plume by interpolating between data from a number of meteorological stations in the region of interest.

The NL sub-system also contains the ISOLA[9] model for very long release durations. This uses statistics of atmospheric conditions and is only appropriate for releases that are sufficiently small that no countermeasures and no early health effects would be expected.

The FL sub-system is linked to the Mesos model[10], developed by Imperial College, UK. This

is a trajectory model for dispersion over long distances that uses meteorological data for a large area, such as the whole of Europe.

Accident consequence assessment programs need to consider the consequences should the accident occur in any of a wide range of atmospheric conditions. It is not possible to calculate the consequences for every sequence of conditions that might arise, and so some method is required to sample a representative set of conditions from those possible. Both the mainframe and PC versions of COSYMA include a flexible program to undertake this sampling.

Only the MUSEMET dispersion model is included in this study, using the NE and NL sub-systems. The uncertainty in dispersion modelling includes both the uncertainty on the spread of the plume around its trajectory, and the uncertainty on the location of the plume trajectory. The other Gaussian models included in COSYMA (RIMPUFF, COSGAP and ISOLA) use similar descriptions of the growth of plumes and of the trajectory. Therefore the uncertainty on consequences predicted using MUSEMET should be similar to the uncertainties predicted using the other Gaussian models. However, MESOS uses a different method of calculating plume trajectories, and the uncertainties on calculations using MESOS may not be the same as those using Gaussian plume or puff models.

B.2 Dose calculations

As stated earlier, COSYMA does not include dosimetric models but uses information from data libraries which are calculated with these models. The libraries include information on the doses from 197 radionuclides.

The data library used for calculating external exposure from γ emitting material deposited on the ground contains outdoor doses per unit deposit integrated to a series of times. These doses are combined with location factors representing the reduction of external γ irradiation by the shielding effects of buildings and typical behaviour of the population. The library is drawn from a number of sources, using results of models developed at NRPB[11,12] and Forschungszentrum für Umwelt und Gesundheit (GSF)[13], Germany. The doses for those radionuclides making major contributions to the dose from fission reactor accidents are derived from a model describing the deposition patterns in urban areas and the subsequent transfer of material between the different surfaces. Location factors are used to describe the protection offered by buildings.

The doses from internal irradiation following ingestion or inhalation are calculated using data libraries of dose per unit intake derived using models which are consistent with those in ICRP publications 56, 67 and 69. COSYMA needs information on the dose received in different periods after the accident, and so this information is included in the data libraries. The method used for calculating doses and risks of health effects in the mainframe version of COSYMA allows for the variation of dose per unit intake with age at intake, and so the libraries contain information on doses for different age groups in the population. The PC version uses a simpler method which only considers the doses to adults.

B.3 Food chain models

COSYMA requires information on the concentration of material in foods as a function of time after the accident. It does not include a food chain model, but uses the results of such models through data libraries which give the activity concentration for a range of radionuclides in a number of foods at a series of times following unit deposition. The concentration of material in foods depends on the time of year at which the deposition occurs. COSYMA uses two data libraries, for deposition in summer and winter. Within a run of COSYMA, the "summer" or "winter" data library is used depending on the date in the year of the meteorological sequence being analysed.

COSYMA uses libraries derived from the NRPB model FARMLAND[14] and the GSF model ECOSYS[15]. The libraries were created using agreed values for the food chain parameters for application within the European Union, but there are differences because of other modelling assumptions made and because of the foods considered in each. The foods which can be considered with FARMLAND are milk, meat and liver from cattle, pork, meat and liver from sheep, green vegetables, grain products, potatoes and other root vegetables. The foods which can be considered with ECOSYS are milk, beef pork, grain products, potatoes and other root vegetables, and leafy and non-leafy green vegetables.

The intakes of these foods are calculated within COSYMA using one of two assumptions about the distribution of food between harvest and consumption. One method assumes that all food consumed is produced locally, and is used in calculating individual ingestion doses. The other method uses information on the amount of food produced in the area of interest, and calculates collective doses on the assumption that all food produced is consumed somewhere.

For this study, the FARMLAND food chain model was used to calculate the uncertainty on concentrations of activity in foods. Doses from ingestion of food were calculated on the assumption that all food consumed is produced locally.

B.4 Countermeasures

COSYMA allows the user to consider the effect of a wide range of countermeasures in reducing the exposure of the population, and gives the user considerable freedom in specifying the criteria at which the actions will be imposed or withdrawn[16].

Sheltering as the only action and sheltering combined with evacuation may be implemented automatically or on the basis of dose. The distribution of iodine tablets, automatically or on the basis of dose, can also be considered. These actions are assumed to be implemented sufficiently rapidly to reduce the risks of both early and late health effects. Relocation is considered as an action to reduce doses and risks over longer time periods. It can be implemented on a dose criterion. Return from evacuation or relocation is also considered on a dose criterion. The effects of decontamination in reducing the period of relocation can be considered. If these actions are initiated on the basis of dose, the user can specify the intervention levels, organs and pathways to be considered, and the time over which the dose is to be integrated. The behaviour of the population considered in the dose criteria can also be described using location factors.

Food restrictions can also be considered[17]. They can be implemented or withdrawn on the basis of doses received within specified time periods or on the basis of the instantaneous concentration of radionuclides in foods.

B.5 Health effects

COSYMA considers both early and late health effects in the population, using methods recommended by NRPB[18,19], the US Nuclear Regulatory Commission[20] and GSF[21].

The risk of early health effects is calculated using "hazard functions". The method allows for the variation of risk with the rate at which dose is accumulated over the first few days following the accident. Ten different fatal and non-fatal effects are considered by COSYMA, though not all are considered for this study.

The risk of late health effects is calculated using the linear dose response relationship. COSYMA considers the risk of fatal and non-fatal cancers in ten organs, and the risk of leukaemia. It also considers the risk of hereditary effects. The method adopted in the mainframe version of COSYMA allows for the variation of risk with age at exposure[22]. PC COSYMA uses a simpler method which only considers the doses and risks to adults, assuming that the risk is the product of committed dose and risk coefficient. The mainframe version of COSYMA can provide information on the numbers of cancers in the people alive at the time of the accident, and in their descendants. It also gives information on the times at which the cancers occur. For this study, the approximation used in PC COSYMA for calculating the risks of late health effects was adopted.

B.6 Economic effects

COSYMA can calculate the off-site economic cost of the accident, considering the costs arising from the countermeasures and the costs of health effects. The assumptions and models are described in references 23 and 24. The countermeasures for which costs are considered are movement of the population, food restrictions and decontamination. The costs arising from lost production in the area from which people are moved can be assessed in terms of the per capita contribution of the relocated population to gross domestic product (GDP) or in terms of the value of the land affected. For longer periods of relocation, the lost capital value of the land and its assets may be calculated. The costs of food restrictions include contributions to GDP as well as the lost capital value and the disposal costs of the food affected. The cost arising from health effects may be calculated in terms of the treatment costs and the lost economic productivity of the affected individuals or an estimation of the cost of health effects may be obtained using a more subjective approach to the valuation of life.

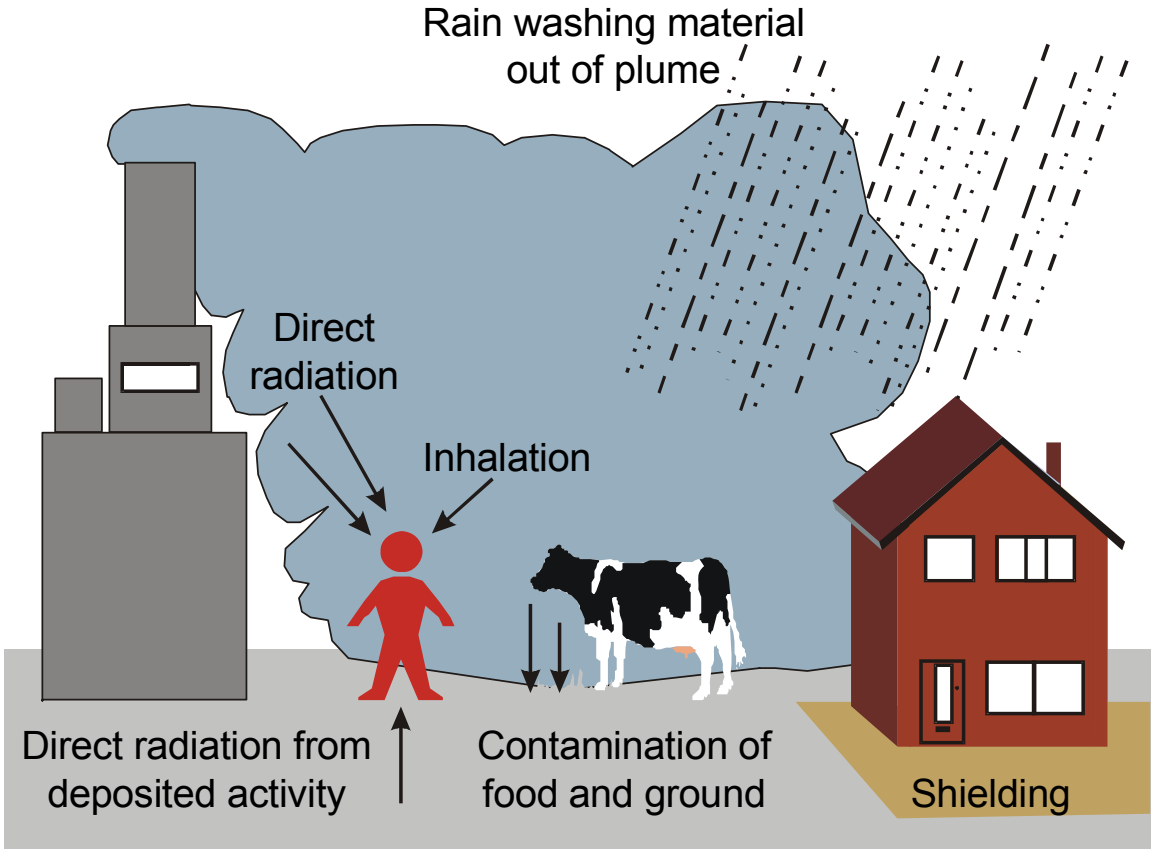
This study did not consider the uncertainty on economic effects.

B.7 References

1 KfK and NRPB. COSYMA: A new program package for accident consequence assessment. CEC. Brussels, EUR-13028 (1991).

2 Hill M D, Simmonds J R and Jones J A. NRPB methodology for assessing the radiological consequences of accidental releases of radionuclides to atmosphere - MARC-1. Chilton, NRPB-R224 (1988) (London HMSO).

3 Ehrhardt J, Burkart K, Hasemann I, Matzerath C, Panitz H-J and Steinhauer C. The program system UFOMOD for assessing the consequences of nuclear accidents. KfK-4330 (1988).
4 Jones J A, Mansfield P A , Haywood S M , Hasemann I, Steinhauer C, Ehrhardt J and Faude D. PC COSYMA (Version 2): an accident consequence assessment package for use on a PC. EUR report 16239 (1995).
5 Panitz H-J, Päsler-Sauer J and Matzerath C. UFOMOD: Atmospheric dispersion and deposition. KfK-4332 (1989).
6 Straka J, Geiβ H and Vogt K J. Diffusion of waste air puffs and plumes under changing weather conditions. Contr. Atmos. Phys. 54 207-221, (1981).
7 Jones J A and Charles D. AD-MARC: The atmospheric dispersion module in the methodology for assessing the radiological consequences of accidental releases. Chilton, NRPB-M72, (1982).
8 Mikkelsen T, Larsen S E and Thykier-Nielsen S. Description of the Riso puff model. Nuclear Technol 76 56-65 (1984).
9 Hübschmann W and Raskob W. ISOLA V: a Fortran-77 code for the calculation of the long-term concentration distribution in the environment of nuclear installations.
10 ApSimon H M and Goddard A J H. Atmospheric transport of radioisotopes and the assessment of population doses on a European scale. CEC Luxembourg EUR-9128 (1983).
11 Charles D, Crick M J, Fell T P and Greenhalgh J R. DOSE-MARC: The dosimetric module in the methodology for assessing the radiological consequences of accidental releases. Chilton NRPB-M74 (1982).
12 Crick M J and Brown J. EXPURT: A model for evaluating exposure from radioactive material deposited in the urban environment. Chilton NRPB-R235 (1990).
13 Jacob P, Paretzke H G, Rosenbaum H, Zankl M. Organ doses from radionuclides on the ground. Part 1: Simple time dependencies. Health Physics 54 617-633 (1988).
14 Brown J and Simmonds J R. FARMLAND a dynamic model for the transfer of radionuclides through terrestrial foodchains. Chilton. NRPB-R273 (1995).
15 Matthies M, Eisfeld K, Müller H, Paretzke H G, Pröhl G and Wirth G. Simulation des Transfers von Radionukliden in landwirtschaflichen Nahrungsketten. GSF Bericht S-882 (1982).
16 Hasemann I and Ehrhardt J. COSYMA: dose models and countermeasures for external exposure and inhalation. Karlsruhe KFK 4333, (1994).
17 Steinhauer C. COSYMA: ingestion pathways and foodbans. Karlsruhe. KfK 4334 (1992)
18 Edwards A A. Private communication (1995).
19 NRPB. Estimates of late radiation risks to the UK population. Documents of the NRPB 4 (4) 15-157 (1993).
20 Evans J S, Moeller D W and Cooper D W. Health effects models for nuclear power plant accident consequence analysis. NUREG/CR-4214 (1985), Rev 1, (1990).
21 Paretzke H G, Stather J W and Muirhead C R. Risk factors for late somatic effects. In Proceedings of the CEC Seminar on methods and codes for assessing the off-site consequences of nuclear accidents, Athens 1990, Luxembourg EUR 13013 (1991).
22 Ehrhardt J, Hasemann I, Matzerath-Boccaccini C, Steinhauer C and Raicevic J. COSYMA: health effects models. Karlsruhe FZKA 5567 (1995).
23 Haywood S M, Robinson C A and Heady C. COCO-1: model for assessing the cost of offsite consequences of accidental releases of radioactivity. Chilton NRPB-R243 (1991).
24 Faude D. COSYMA: Modelling of economic consequences. Karlsruhe, KfK Report 4336 (1992)



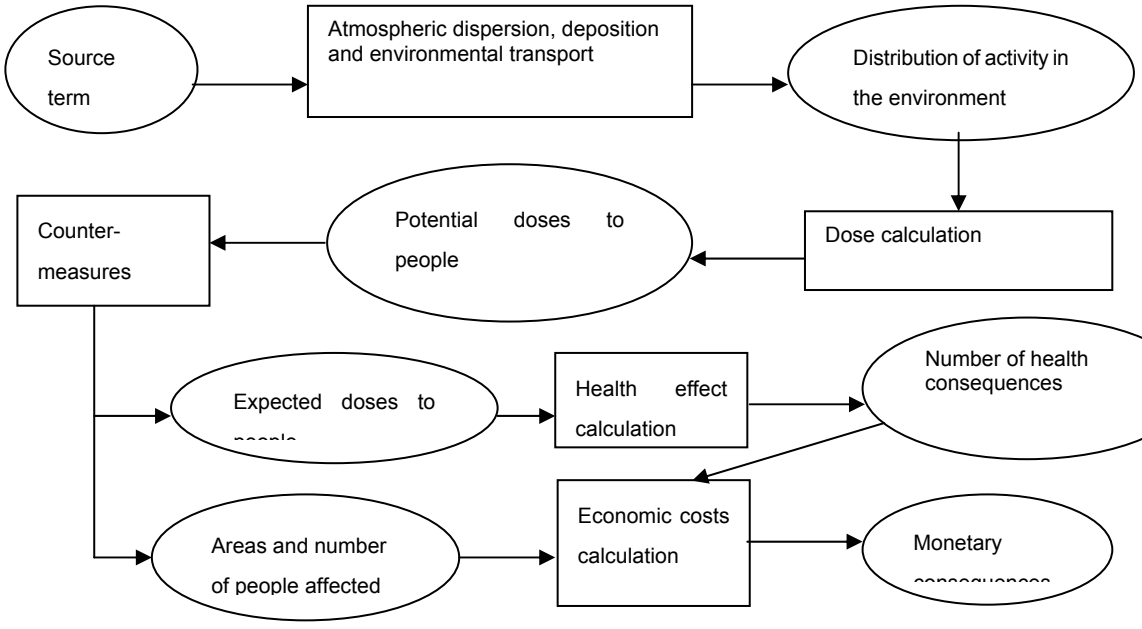


Figure B.1 Processes modelled in COSYMA

APPENDIX C

Processing Expert Judgements

In this chapter we introduce two mathematical techniques which enable us to translate the information given by the experts to information which is required in doing the uncertainty analysis. The first mathematical technique is termed *Probabilistic Inversion*. This technique translates information on the output of a model to the input of the model, in such a way that if the distribution on the input of the model is propagated through the model it resembles as closely as possible the information available. The second mathematical technique is concerned with transforming conditional probabilities of multiple experts on an event into *one* rank correlation.

PROBABILISTIC INVERSION

Suppose we had a distribution over code input parameters of a particular model. We could then push this distribution through the model and obtain a distribution over its output. The problem at hand involves reversing this procedure: we have quantiles of the distributions of the model output given by experts and we seek a distribution over the code input parameters which, when pushed through the model, yields quantiles over the model output agreeing with those from the experts. Hence our problem is one of probabilistic inversion: we must invert the model so as to 'pull back' the distribution over the model output onto the parameters of the model.

Let H represent a distribution over model output. Let F represent a distribution over the input parameters of the model, and let $G(F)$ represent the distribution over model output, obtained by pushing the distribution F through the model.G. Then our problem may be represented as: Find F such that $G(F) \approx H$, where '$\approx$' means 'has the same distribution as', or equivalently, $F \approx G^{-1}(H)$.

Note that a probabilistic inverse $G^{-1}(H)$ may not exist, and if it exists it will in general not be unique. Therefore we must have a method of selecting a preferred distribution in case of non-uniqueness and a method of choosing a best fitting distribution in case of non existence.

Note also that probabilistic inversion is not restricted to expert judgment only. Distributions obtained from a series of experiments under similar conditions can also be used as input in the probabilistic inversion technique. For a detailed description of probagbilistic inversion, see reference 1.

Target Variables & Elicitation Variables

In performing uncertainty analysis, a distribution on the code input parameters must be specified. The code input parameters will be termed target variables. Variables which are physically observable and for which the experts provide information will be called elicitation variables. The models in the code are idealizations to which the experts may not subscribe; moreover, the code input parameters may not correspond to physical measurements with which the experts are familiar. An important element of the methodology adopted in the joint study is to

query experts on distributions on potentially measurable quantities. Target variables cannot be always elicitation variables; elicitation variables will be constructed which are related to the target variables and which represent physically measurable quantities with which the experts are familiar.

Let us illustrate the distinction between target and elicitation variables by looking at three examples from the joint project:

Example 1 In some cases the code input parameters correspond to physically measurable quantities with which the experts are familiar. For example, deposition velocities to various surfaces under various conditions are directly measurable. The measured values are known to depend on a large number of physical parameters which cannot be measured or controlled on any given experiment. Moreover, the functional form of the dependence is not known. Hence, if a controlled experiment is repeated many times, different values will be found reflecting different values of uncontrolled and unknown physical parameters. If a measurement set-up is described to an expert, (s)he can express his/her uncertainty via a subjective distribution over possible outcomes of the measurement. In such cases the experts are questioned directly about the uncertainty with respect to code input parameters.

Example 2 The lateral plume spread σ_y is modelled as a power law function of downwind distance x from the source:

$$\sigma_y(x) = P_y x^{Q_y} \qquad (1)$$

where the dispersion coefficients P_y and Q_y depend on the stability of the atmosphere at the time of the release. (1) is not derived from underlying physical laws, rather, the coefficients are fit to data from tracer experiments. For the uncertainty analysis of COSYMA, we require distributions on P_y and Q_y which, when pushed through (1), will yield the uncertainty on σ_y for each downwind distance x . P_y and Q_y are the target variables. Although the experts have experience with measured values of σ_y under various conditions, it is unrealistic to expect them to be able to quantify their uncertainty in terms of the target variables P_y and Q_y. Indeed, the dimension of P_y must be $[\text{meters}]^{1-Q_y}$. In this case, elicitation variables Y_i were defined on the lateral plume spread for downwind distances $x_1, \ldots, x_n$ and uncertainty distributions were queried on these quantities. Note that the elicitation variables will be represented by Y_i to indicate that the experts were not restricted in using model (1) in making their assessments. The problem then arises how to translate these elicited distributions into distributions on the target variables P_y and Q_y.

Example 3 The migration of radioactive material through various depths of soil is modelled using a so-called compartmental model, see Figure C.1.

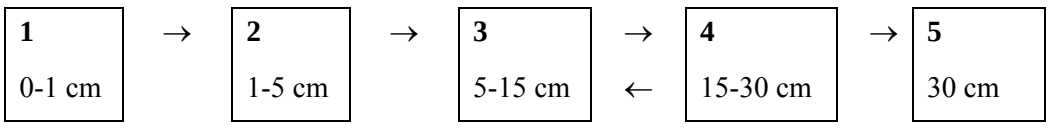


Figure C.1: Compartmental model of soil-migration

The target variables for the code are transfer coefficients k_{ij}, which represent the proportion of material moved from box *i* to box *j* in a small time interval. Based on Figure C.1, a set of first order differential equations can be constructed which, with the appropriate initial conditions, fully specifies the movement of the material between the boxes. The aim is to derive a distribution on the transfer coefficients. Transfer coefficients cannot be measured directly and therefore cannot be elicitation variables. In this case the query variables were on times T_j when half of the mass of the deposited material has past beyond box *j*. From this information a distribution on the transfer coefficients has to be determined.

The determination of a distribution on the various target variables (P_y, Q_y or transfer coefficients), given information on query variables (Y_i *or* T_j) is an example of probabilistic inversion.

In total, information on 23 models were probabilistically inverted:

Dispersion & deposition : Gaussian model for 4 stability classes and wet deposition for methyl-iodide, elemental iodine and aerosol particles.

Foodchain : Soil migration model for 2 nuclides, grain model for 2 nuclides and rootcrop model for 2 nuclides.

Early Health Effects : mortality for GI-tract, lung and whole body and morbidity for skin and lung.

Internal Dosimetry : lung model, absorption to the blood for 7 different nuclides, Systemic retention for 7 different nuclides.

In the next section, two solution schemes for probabilistic inversion will be discussed.

Solution Schemes

For a detailed description of the solution schemes see references 2 - 4. The experts were queried on the 5%, 50% an 95% quantile points of their distribution. The space where the target variables are defined will be termed target variable space and the space where the elicitation variables are defined will be termed observable space. First the probabilistic inversion technique implemented in PARFUM is introduced, followed by a more advanced probabilistic inversion technique, which has been implemented in PREJUDICE. This section is concluded with a discussion on the differences among the various solution schemes.

We will illustrate the different steps of PARFUM and PREJUDICE using Example 2, the lateral plume spread. In performing the uncertainty analysis, a distribution over (P_y, Q_y) is required. As discussed above, it was concluded that the dispersion coefficients cannot serve as elicitation variables. Therefore the elicitation variables Y_i are on the lateral plume spread for downwind distances $x_1, \ldots, x_n$. The different figures illustrate the steps of PARFUM and PREJUDICE in case of $n = 2$.

PARFUM

The acronym PARFUM stands for PARameter Fitting for Uncertain Models.

Step 1 Support of Distribution: In this step the support of the distribution over the target variables is determined.
For $j \in \{1,2,3\}$ each $i, i = 1,\ldots,n$, the j – th quantile point is chosen for Y_i, say $y_{I,j}$, where $j \in \{1,2,3\}$ and $j = 1$ correspond to the 5% quantile, $j = 2$ to the 50% quantile and $j = 3$ to the 95% quantile. The set

$$s = \left(y_{1,j}, \ldots, y_{n,j}\right) \tag{2}$$

is called a scenario. Let S represent the set of all scenarios. Each scenario is tested for physically admissibility; in this example $s \in S$ is admissible only if for all $j : y_{i,j} < y_{i+l,j}$ where $i = 1,\ldots,n - 1$. Let S^* denote the set of admissible scenarios.

For each scenario $s \in S^*$, estimates (p_y^s, q_y^s) for the target variables are determined, such that the sum of squared errors between model output and scenario s is minimized.

$$\sum_{i=1}^{n} (p_y^s x_i^{q_y^s} - s_i)^2 \tag{3}$$

For each target variable, an interval is determined which covers all estimates. A uniform grid of points from the intervals for P_y and Q_y is propagated through the power law model (1). The grid points taken in the target variable space form the support of the distribution over the target variables and will be indicated by I. PARFUM generates n observable spaces of dimension 1. See Figure C.2 for a graphical illustration, the pair (p_y, q_y) is sampled uniformly and mapped into the observables ($\sigma_y(x_1)$ and $\sigma_y(x_2)$).

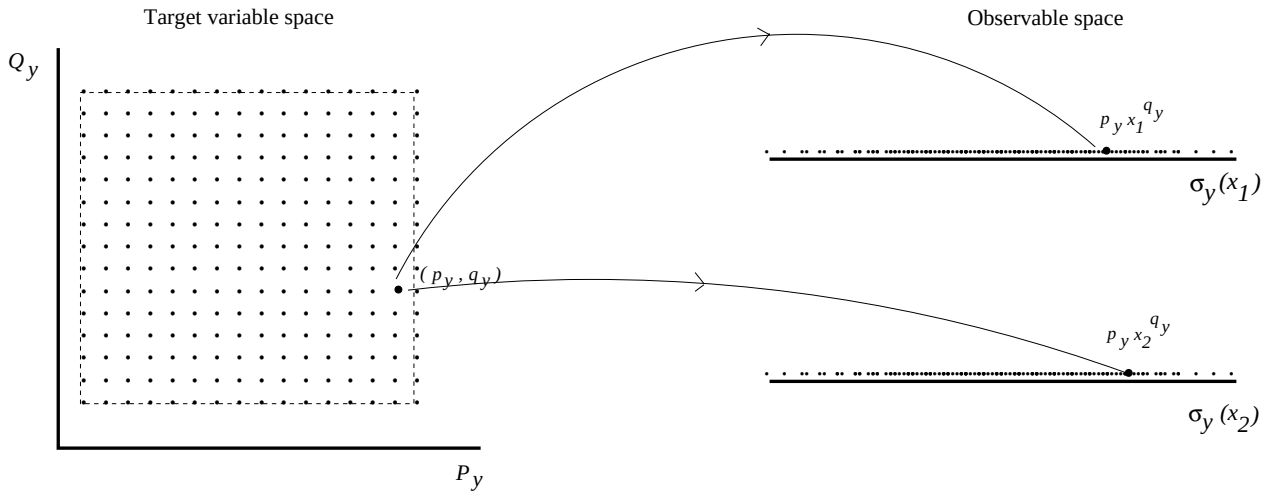


Figure C.2: PARFUM: Propagation of samples

Step 2 Determination of Distribution: for each observable $\sigma_y(x_i)$ a distribution P_{Y_i} over the propagated samples is determined. These distributions have minimum relative information with

respect to the uniform background measure and their quantile information complies with the quantile information of the distribution of Y_i as given by the experts. As a result, propagated samples contained in the same interquantile interval receive the same probability; for example the propagated samples contained in interquantile interval ($y_{I,50\%}$,$y_{I,95\%}$] for elicitation variable Y_i receive probability:

$$P_{Y_i,3} = \frac{0.45}{\sum_{(p_y,q_y)\in I} 1_{\{(p_y,q_y)|p_y x_i^{q_y} \in (y_{i,50\%},\, y_{i,95\%}]\}}(p_y,q_y)}$$

where $1_A(x) = 1$ if $x \in A$ and zero otherwise. As each point in I can be associated with a propagated sample, we have n distributions P_i on I, see Figure C.3. A single distribution P on I is required, based on these n distributions P_i. PARFUM determines the distribution P such that $\sum_{i=1}^{n} I(P_i|P)$ is minimum where $I(P_I\backslash P)$ is the relative information of P_i with respect to P, for details see [8].

PREJUDICE

The acronym PREJUDICE stands for PRocessing Expert JUDgment Into Code paramEters. Again, the different steps will be illustrated using Example 2, the lateral plume spread.

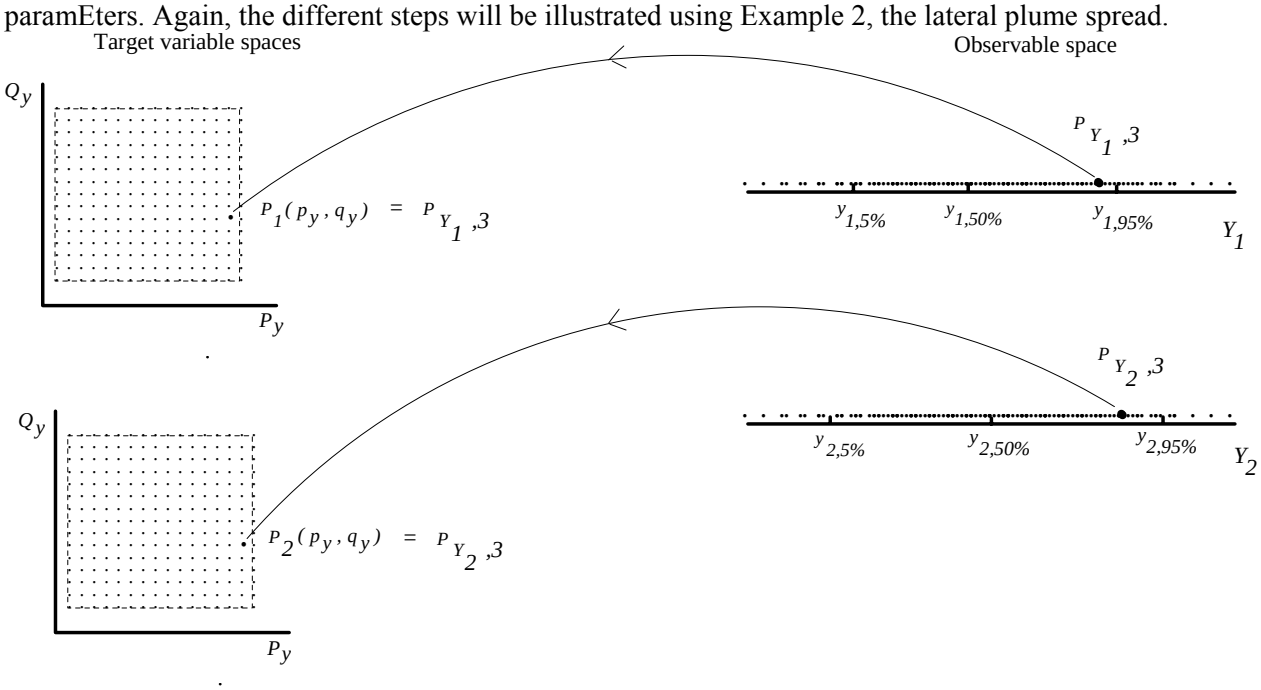


Figure C.3 PARFUM: 'Pullback'-distributions

Step 1 Support of Distribution: In this step the support of the distribution over the target variables is determined.

For each i, $i = 1,\ldots,n$, an elicited quantile is chosen for Y_i, say y_{i,j_i} where $j_i \in \{1,2,3\}$. The set

$$s = (y_{1,j_1},\ldots,y_{n,j_n}) \qquad (4)$$

is called a scenario. Let S be the set of all scenarios. Next, each scenario is tested for physically admissibility; in this example $s \in S$ is admissible only if $y_{i,j_i} < y_{i+1,j_{i+1}}$, where $i = 1,...,n-1$. Let S^* denote the set of admissible scenarios.

For each scenario $s \in S^*$, estimates (p_y^s, q_y^s) for the target variables are determined, such that the sum of squared errors between model output and scenario s is minimized

$$\sum_{i=1}^{n} (p_y^s x_i^{q_y^s} - s_i)^2 \qquad (5)$$

For each target variable, intervals are determined which are (1) mutually exclusive and (2) whose union covers all estimates. In this way, each estimate (p_y^s, q_y^s) is contained in one hypercube of the target variable space. A number of samples are taken uniformly from each such hypercube and propagated through the power law model (1) for $x = x_1,\ldots,x_n$.. This generates a set of points in the observable space. The samples taken in the target variable space will be the support of the distribution over the target variables and will be indicated by I. See Figure C.4 for a graphical illustration, the pair (p_y, q_y) is sampled uniformly from the rectangle $B_{1,5} \times B_{2,2}$ and mapped into the observable pair $(\sigma_y(x_1), \sigma_y(x_2))$.

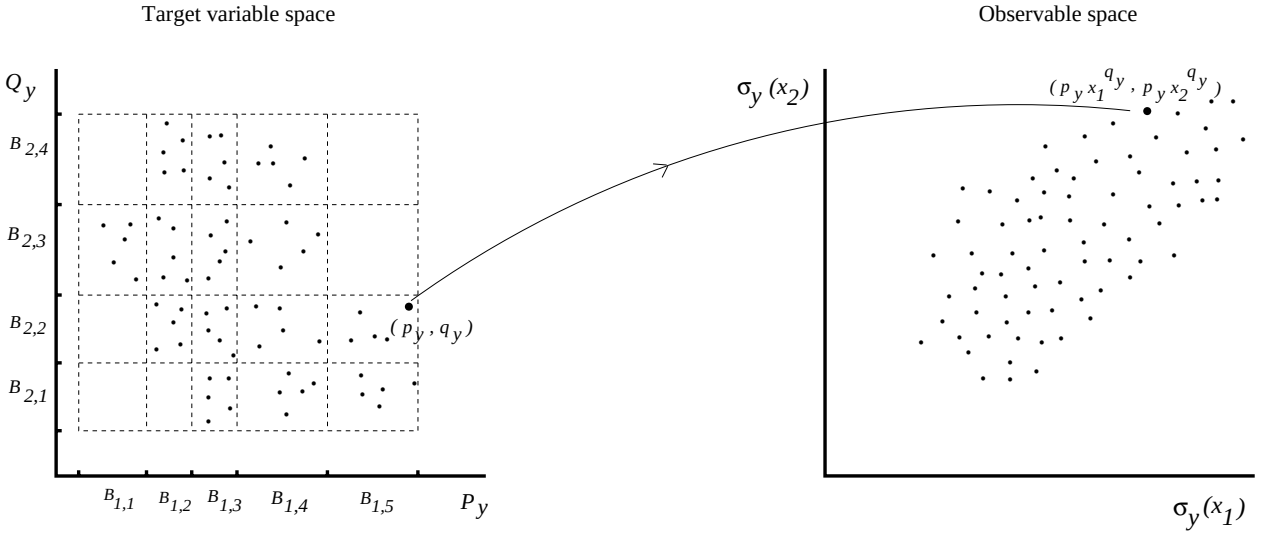


Figure C.4: PREJUDICE: Propagation of samples

Step 2 Determination of Distribution: Note that the axes of the observable space can be associated with elicitation variables $Y_i, i = 1,...,n$. The 5%, 50% and 95% quantiles of the distributions of each elicitation variables are available.

Briefly, a distribution over the propagated samples in the observable space is determined, which has minimum relative information with respect to the uniform background measure and such that for each elicitation variableY_i, the quantile information of its respective marginal distribution complies with the quantile information of the distribution of Y_i.

Each propagated sample in the observable space thus receives a probability and each point in I is assigned the probability associated with its image in the observable space. In this way a distribution over I is determined. Details are given in the next section.

PREJUDICE: Computation of Distribution

Based on the contraints on the quantile information on Y_i, we introduce a set of hypercubes in the observable space. For Y_i , we distinguish the four interquantile intervals

$$\left(-\infty, y_{i,5\%}\right], \left(y_{i,5\%}, y_{i,50\%}\right], \left(y_{i,50\%}, y_{i,95\%}\right], \left(y_{i,95\%}, \infty\right)$$

Taking the product of all such intervals, for all elicitation variables, we generate a set of "observable hypercubes", indexed as $i_1 \ldots i_n$, where $i_j \in \{1,2,3,4\}$. Thus $i_j = 3$ means that we consider interval $(y_{j,50\%}, y_{j,95\%}]$ for elicitation variable Y_j. For 2 elicitation variables, Y_1 and Y_2 the observable hypercubes are shown in Figure C.5.

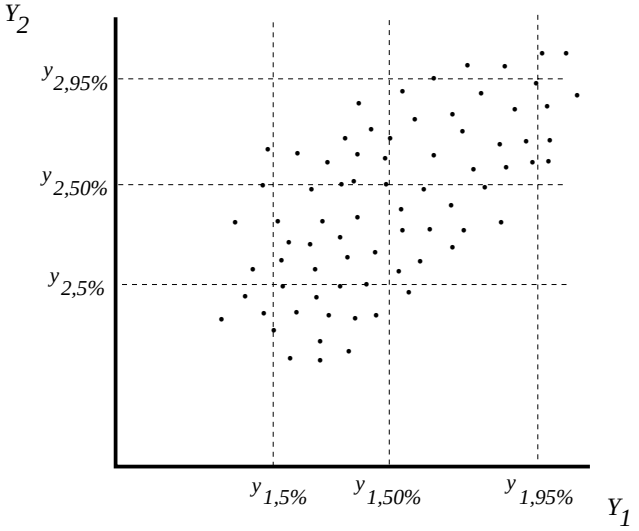


Figure C.5: PREJUDICE: Observable Hypercubes

The number of distributions over the propagated samples which will satisfy the constraints on the quantile information on Y_i, as described in Step 2, may be large. From this set of distributions we want to select one distribution: the distribution which has minimum relative information with respect to the uniform background measure[5,6]. Determining the distribution which has minimum relative information with respect to the uniform background measure can be formulated as a Convex Programming (CP) problem.

Suppose K samples from the target variable space are propagated to the observable space. Let $k (k = 1,\ldots K)$ indicate the k -th sample taken in the target variable space, which is propagated through the Equation (1), evaluated at $x_1,\ldots,x_n$, to generate $\sigma_y^k(x_1),\ldots,\sigma_y^k(x_n)$. Furthermore let p_k represent the probability of the k -th sample. The NLP problem may be formulated as follows:

$$min \sum_{k=1}^{K} p_k \log \frac{p_k}{C} \tag{6}$$

$$\sum_{k=1}^{K} p_k 1_{\left[\sigma_y^k(x_1) \le y_{1,5\%}\right]}(k) = 0.05 \qquad \ldots \qquad \sum_{k=1}^{K} p_k 1_{\left[\sigma_y^k(x_n) \le y_{n,5\%}\right]}(k) = 0.05$$

$$\sum_{k=1}^{K} p_k 1_{\left[y_{1,5\%} < \sigma_y^k(x_1) \le y_{1,50\%}\right]}(k) = 0.45 \qquad \cdots \qquad \sum_{k=1}^{K} p_k 1_{\left[y_{n,5\%} < \sigma_y^k(x_n) \le y_{n,50\%}\right]}(k) = 0.45$$

$$\sum_{k=1}^{K} p_k 1_{\left[y_{1,50\%} < \sigma_y^k(x_1) \le y_{1,95\%}\right]}(k) = 0.45 \qquad \cdots \qquad \sum_{k=1}^{K} p_k 1_{\left[y_{n,50\%} < \sigma_y^k(x_n) \le y_{n,95\%}\right]}(k) = 0.45$$

$$\sum_{k=1}^{K} p_k 1_{\left[\sigma_y^k(x_1) > y_{1,95\%}\right]}(k) = 0.05 \qquad \cdots \qquad \sum_{k=1}^{K} p_k 1_{\left[\sigma_y^k(x_n) > y_{n,95\%}\right]}(k) = 0.05$$

$$p_k \ge 0$$

where $1_A(x) = 1$ if $x \in A$ and zero otherwise and $C = \frac{1}{K}$. Based on the Karush-Kuhn-Tucker conditions for CP problem (6), it can be shown that if $\sigma_y^k(x_1), \ldots, \sigma_y^k(x_1)$ and $\sigma_y^l(x_1), \ldots, \sigma_y^l(x_1)$ fall in the same observable hypercube, then $p_k = p_l$. This observation allows us to formulate the NLP problem in terms of observable hypercubes:

$$min \sum_{i_1=1}^{4}\sum_{i_2=1}^{4}\cdots\sum_{i_n=1}^{4} c_{i_1\ldots i_n} p^*_{i_1\ldots i_n} \log p^*_{i_1\ldots i_n} + C^* \qquad (7)$$

$$\sum_{i_2=1}^{4}\cdots\sum_{i_n=1}^{4} c_{1i_2\ldots i_n} p^*_{1i_2\ldots i_n} = 0.05 \qquad \cdots \qquad \sum_{i_1=1}^{4}\cdots\sum_{i_{n-1}=1}^{4} c_{i_2\ldots i_{n-1}1} p^*_{i_1\ldots i_{n-1}1} = 0.05$$

$$\sum_{i_2=1}^{4}\cdots\sum_{i_n=1}^{4} c_{2i_2\ldots i_n} p^*_{2i_2\ldots i_n} = 0.45 \qquad \cdots \qquad \sum_{i_1=1}^{4}\cdots\sum_{i_{n-1}=1}^{4} c_{i_2\ldots i_{n-1}2} p^*_{i_1\ldots i_{n-1}2} = 0.45$$

$$\sum_{i_2=1}^{4}\cdots\sum_{i_n=1}^{4} c_{3i_2\ldots i_n} p^*_{3i_2\ldots i_n} = 0.45 \qquad \cdots \qquad \sum_{i_1=1}^{4}\cdots\sum_{i_{n-1}=1}^{4} c_{i_2\ldots i_{n-1}3} p^*_{i_1\ldots i_{n-1}3} = 0.45$$

$$\sum_{i_2=1}^{4}\cdots\sum_{i_n=1}^{4} c_{4i_2\ldots i_n} p^*_{4i_2\ldots i_n} = 0.05 \qquad \cdots \qquad \sum_{i_1=1}^{4}\cdots\sum_{i_{n-1}=1}^{4} c_{i_2\ldots i_{n-1}4} p^*_{i_1\ldots i_{n-1}4} = 0.05$$

$$p^*_{i_1\ldots i_n} \ge 0$$

where $c_{i_1\ldots i_n}$ represents the number of propagated samples falling in observable hypercube $i_1\ldots i_n$, and $C^* = \log\left(\sum_{i_1=1}^{4}\cdots\sum_{i_n=1}^{4} c_{i_1\ldots i_n}\right)$. Furthermore, note that $\sum_{k=1}^{K} p_k = 1$ but in general $\sum_{i_1=1}^{4}\cdots\sum_{i_n=1}^{4} p^*_{i_1\ldots i_n} \ne 1$. CP problem (7) is solved for $p^*_{i_1\ldots i_n}$. The $p^*_{i_1\ldots i_n}$ represent a value of each sample in observable hypercube $i_l\ldots i_n$. The interior point solver for large-scale convex problems [1] avoids problems with $p^*_{i_1\ldots i_n} = 0$ and gave excellent performance. In Figure C.6 the highlighted point receives the value $p^*_{3,4}$ and this value is assigned also to the pre-image of this point, (p_y, q_y) in the target variable space.

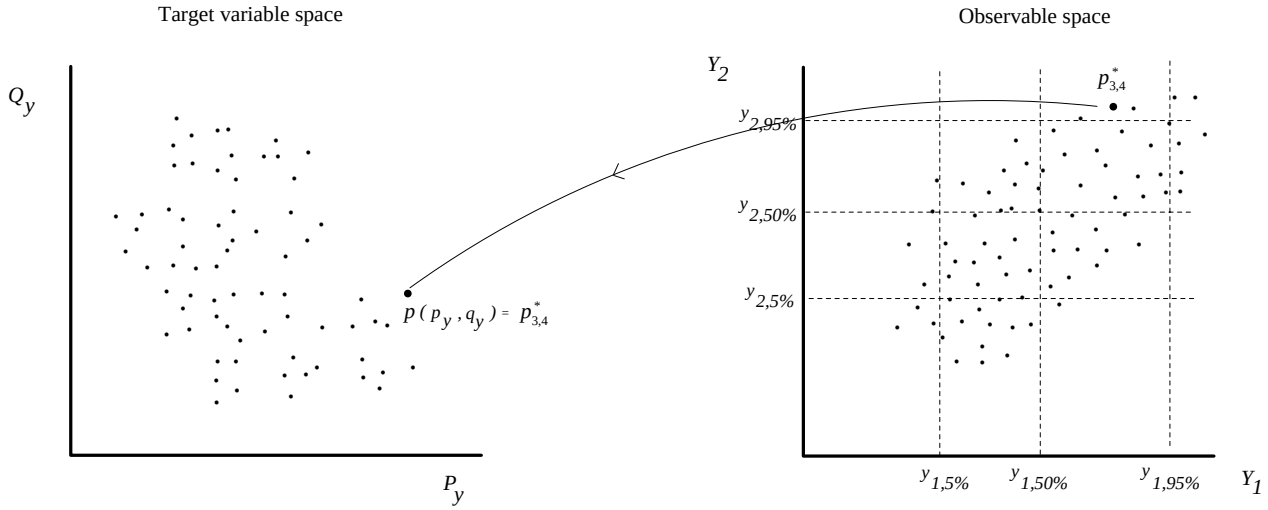


Figure C.6: PREJUDICE: 'Pullback'-distribution

Dealing with infeasibilities

The CP problem (7) may not be feasible. In this case we will reduce the dimension of the observable space. Still considering Example 2 with *n* elicitation variables, suppose the CP problem (7) is infeasible. The idea is to look at $\frac{n!}{(n-1)!}$ problems of dimension *n-1* . For each of the $\frac{n!}{(n-1)!}$ problems, **Step 1** is carried out. Next let *I* be the union of $\frac{n!}{(n-1)!}$ search grids.

Generate $\frac{n!}{(n-1)!}$ CP problems and let $N\left(N \le \frac{n!}{(n-1)!}\right)$ denote the number of CP problems for which a distribution in **Step 2** can be determined. Assuming that *N>0*, we have obtained *N* distributions over the target variables on their specific supports. We are now confronted with the problem of finding a distribution over the target variables which "best fits" these *N* distributions. For the solution of this problem we refer to [8].
If *N=0* we will reduce the dimension of the problem once more, and perform **Step 1** and **Step 2** for $\frac{n!}{2(n-2)!}$ problems of dimension *n-2*.

PARFUM vs. PREJUDICE

Probabilistic inversion means finding a distribution over code input parameters which yields distributions over model predictors closely matching elicited distributions. In the joint study the elicited distributions are obtained by structured expert judgement, but of course other sources of distributions would serve just as well.

The PARFUM method is easily implemented and always yields a feasible problem. For small problems the numerical results have been acceptable, but not really good. PARFUM is conceptually flawed in that it optimizes a citerion which is not really the criterion of interest in judging performance. PREJUDICE repairs this conceptual flaw and yields excellent performance.

Both methods are strongly driven by heuristics (**Step 1** for both methods) in choosing a set of initial code input parameters over which an optimal distribution is sought.

DEPENDENCIES

It has long been known that significant errors in uncertainty analysis can be caused by ignoring dependencies between uncertainties[7,8]. In the joint project, dependency information is obtained from 3 different sources

1. Rank correlation coefficients extracted from the distribution among target variables obtained via Probabilistic inversion.
2. Dependencies among elicitation variables extracted from the experts.
3. Dependencies specified by project staff.

Given a joint distribution, it is straightforward to determine its corresponding rank correlation matrix. Therefore we will not elaborate on dependencies obtained from 1. As the technique to extract/specify dependencies from experts/project staff is the same, we will only focus on how dependencies were extracted from experts.

In order to obtain dependency information from the experts, new techniques for estimating and analyzing have been developed in the course of the joint effort. We will introduce the procedure how dependencies were elicited from experts, and how the dependencies of multiple experts are combined.

The method of assessing dependencies as described in this section has been used throughout all expert panels of the project, except for the dispersion and deposition panel.

Elicitation of dependencies

The best source of information about dependencies is often the experts themselves. The most thorough approach would be to elicit directly the expert's joint distributions. The practical drawbacks to this approach have forced analysts to look for other dependence elicitation strategies. One obvious strategy is to ask experts directly to assess a (rank) correlation coefficient. Even trained statisticians have difficulty with this type of assessment task[9].

For this reason a new method was developed for eliciting dependencies from experts, which had to be easily understandable to the experts. Suppose we want to elicit the dependence between (continuous) elicitation variables X and Y; firstly the experts assess the marginal distributions for both X and Y, next they are asked to give their conditional probability to the following question:

Suppose Y were observed in a given experiment and its value were found to lie above the median value for Y ; what's your probability that, in this same experiment X would also lie above its median value?

Briefly, a conditional probability of 0 follows if *X* and *Y* are rank correlated by -1, a conditional probability of 0.5 follows if *X* and *Y* are uncorrelated and a conditional probability of 1 emerges if *X* and *Y* are rank correlated by 1. Experts quickly became comfortable with this assessment technique, acknowledged its importance and provided answers which were meaningful to the project staff.

In every expert panel, project staff identified a large number of potential dependencies. For two reasons it was decided that not all dependencies were to be elicited; firstly there are too many questions and secondly, in eliciting all dependencies, it is almost impossible to ensure that the resulting correlation matrix is positive definite. Therefore it was decided that the experts were to be elicited on a selection of all possible dependencies. If this selection is such that the resulting dependency graph is acyclic∗, it is possible to find a joint distribution which

- satisfies the marginal distribution of the selected variables
- has a rank correlation matrix which is positive definite‡ and satisfies the results as specified in the dependency structure.

Moreover, there is a unique completion which has minimal information with respect to the (minimally informative) independent distribution. It is for this reason that the correlations specified by experts are restricted to a tree. A simple algorithm may be employed to find the

---

∗ An undirected graph is a set of nodes, N and a set of edges, E, where an edge is an unordered pair of nodes. An undirected graph is acyclic if its edge set contains no cycles, i.e. there is no e1,...en in E such that $e_1 = \{a_1,b_1\}$, $e_2 = \{b_1,b_2\}$,...$e_n = \{b_{n-1},a_1)$. An undirected acyclic graph is called a tree.

‡ If $a = (a_1,..a_N)^t$ denotes a column vector (^t denotes transposition), then $a^ta = a_1^2+...+a_N^2$. By definition, an N x N matrix X is positive definite if for any non zero vector a, $a^tXa > 0$. This is a generalization of the property that the square of a number must be positive. Indeed, it can be shown that any positive definite matrix X can be written as Y^tY with Y triangular (the Cholesky decomposition) hence the condition becomes

$0 < a^tXa = a^tY^tYa = (Ya)^tYa$.

If $X = X_1,..X_N$ is a vector of standardized random variables then the covariance matrix is

E(X^tX), (note, X^tX is an N x N matrix).

Alternatively, replace X_i by a column vector of M realizations of X_i. Then $X = \{X_{ij}\}$ is an M x N matrix, X^tX is an N X N matrix whose ij-th element is $X_{1i}X_{1j} + X_{2i}X_{2j},...+ X_{Mi}X_{Mj}$, which is the empirical expectation of the product of random variables X_i and X_j. Since X_i and X_j are standardized, this is also the Pearson correlation coefficient of X_i and X_j. This correlation matrix is positive definite, since for any nonzero N-vector a

$a^tX^tXa = (Xa)^tXa > 0$.

If a correlation matrix is partially specified, then the problem whether this can be extended to a positive definite matrix is known as the matrix completion problem, and in general is quite difficult. If the specified cells cannot be extended to a positive definite matrix, then the specification is inconsistent: there is no set of random variables with the specified correlations. Of course, since a correlation matrix is necessarily symmetric, it suffices to specify only the upper triangular part (the entries on the main diagonal are 1's). If the specified cells of the correlation matrix constitute the edges of a tree, then if these specified cells are between -1 and 1, the matrix can always be extended to a positive definite matrix, and hence the specification is always consistent.

minimally informative completion.

From the set of distributions, which share the properties as stated above, we select the distribution which has minimum relative information (with respect to the product distribution) among all the distributions[5].

The list of dependency questions is constructed by members of the project staff, the consequence analyst and the uncertainty analyst. The consequence analyst drafts a list with potential important dependencies among elicitation variables. This list is then reviewed by the uncertainty analyst to see if the corresponding dependency structure would result in an acyclic dependency structure. After a number of iterations the final dependency document is given to the experts.

Combining of Conditional Probabilities

Expert *e* assesses $\pi^e_{r_1,r_2}(X,Y) = P\left(F^e_X(X) > r_1 \middle| F^e_Y(Y) > r_2\right)$ for $r_1 = r_2 = \frac{1}{2}$. Furthermore, it is assumed that F^e_X and F^e_Y are expert *e*'s (continuous invertible) cumulative distribution functions (cdf's) for *X* and *Y* respectively. $F^e_X(X)$ is called expert *e*'s quantile function of *X* and which is uniformly distributed on the interval [0,1].

Having elicited the conditional probabilities $\pi^e_{\frac{1}{2},\frac{1}{2}}(X,Y)$ from the experts, the next step is to combine them into one conditional probability. In general the medians of the experts will be different, for this reason one cannot combine the conditional probabilities $\pi^e_{\frac{1}{2},\frac{1}{2}}(X,Y)$ via linear pooling; the pooling will not be over the same event.

Briefly, the marginal distributions of the elicitation variables *X* and *Y* of the experts can be pooled, resulting in cdf's F^{DM}_X and F^{DM}_Y for *X* and *Y* for the Decision Maker (DM). Let $x_{DM,50}$ and $y_{DM,50}$ and denote the medians for the DM's distribution for *X* and *Y*. A minimum relative information distribution is associated with each expert *e*. We can compute the conditional probabilities $P_e(X > x_{DM,50} | Y > y_{DM,50})$. Since these conditional probabilities are defined over the same events for all experts, they can be combined via the linear pool. This yields a value for $\pi^{DM}_{\frac{1}{2},\frac{1}{2}}(X,Y)$ for DM, from which we can find the corresponding Spearman's rank correlation $\rho^{DM}_{X,Y}$.

Spearman's rank correlation $\rho_{X,Y}$ and conditional probability $\pi_{\frac{1}{2},\frac{1}{2}}(X,Y)$

Consider all joint distributions for *(X,Y)* having marginals F_X, F_Y, having minimum information relative to the distribution with independent marginals and having Spearman's rank correlation $\rho_{X,Y} \in [-1,1]$. For each $\rho_{X,Y} \in [-1,1]$ there is a unique value for $\pi_{\frac{1}{2},\frac{1}{2}}(X,Y) \in [0,1]$. For this distribution we compute the values $\pi_{r_1,r_2}(X,Y)$. Figure 7 shows values for $\pi_{r_1,r_2}(X,Y)$, where r_1 and r_2 ranges for $\rho_{X,Y}$ = -0.9,-0.8,...,0.8,0.9.

Hence, we may consider the $\rho_{X,Y}$ characterizing the minimal information distribution as a function of $\pi_{\frac{1}{2},\frac{1}{2}}(X,Y)$.

The unique relationship between $\rho_{X,Y}$ and $\pi_{\frac{1}{2},\frac{1}{2}}(X,Y)$ enables us to build a table which links exceedence probabilities $P(F_X(X) > r_1)$ and $P(F_Y(Y) > r_2)$, Spearman's rank correlation $\rho_{X,Y}$ and conditional probability $\pi_{\frac{1}{2},\frac{1}{2}}(X,Y)$. This table has been generated using the simulation program UNICORN[6]. The values of r_1 and r_2 were taken from the set {0.0025, 0.05, 0.1, 0.2,..., 0.8, 0.9, 0.95, 0.9975}. The rank correlations $\rho_{X,Y}$ were taken from the set {-1,-0.9,...,0.9,1} .

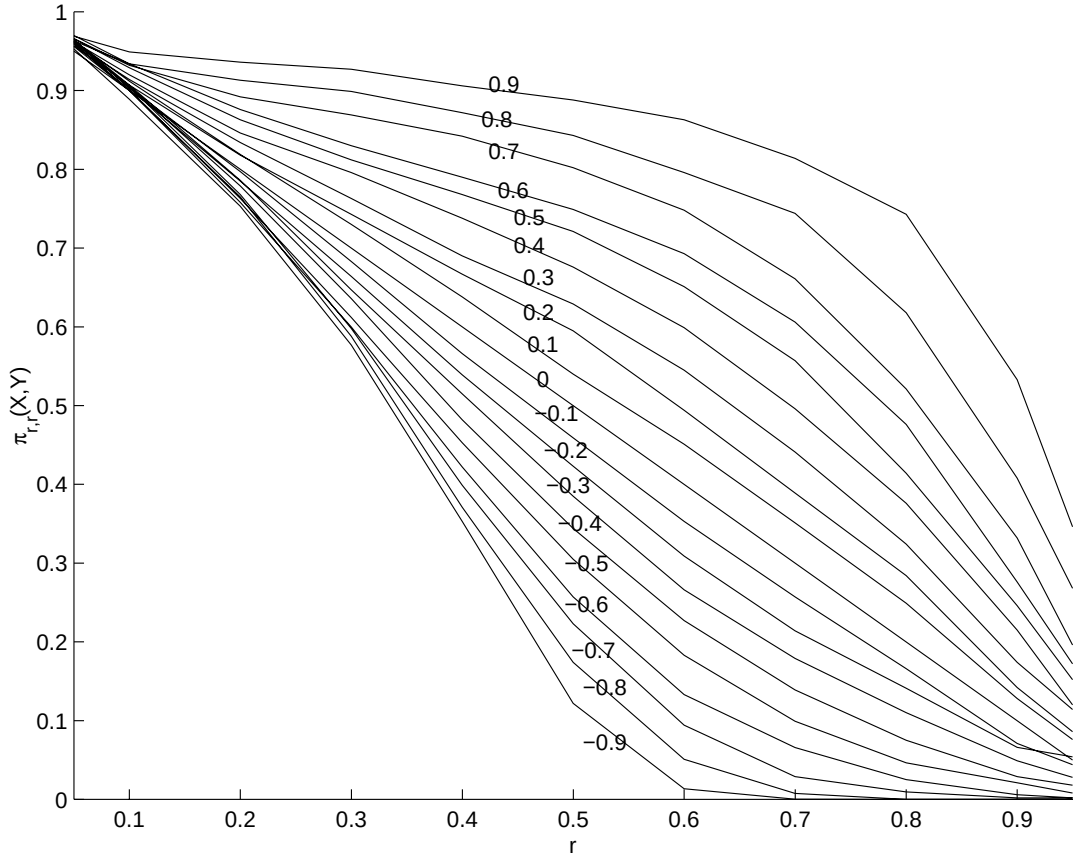


Figure C.7: Relationship between $\pi_{r,r}(X,Y)$ and r for different Spearman's rank correlations $\rho_{X,Y}$.

Solution scheme

The technique of combining conditional probabilities into one rank correlation can be summarised by 6 steps:

Step 1 : For each expert e query $\pi^e_{\frac{1}{2},\frac{1}{2}}(X,Y) = P\left(F^e_X(X) > \frac{1}{2} \middle| F^e_Y(Y) > \frac{1}{2}\right)$

Step 2 : For each expert e, find $\rho^e_{X,Y}$ from the table which passes through $\left(\frac{1}{2},\frac{1}{2},\pi^e_{\frac{1}{2},\frac{1}{2}}(X,Y)\right)$

Step 3 : Take linear pooling of experts' marginals to determine F_X^{DM} and F_Y^{DM} , find $x_{DM,50}$, $y_{DM,50}$

Step 4 : For each expert *e*;

• Determine $P_e(X > x_{DM,50})$ and $P_e(Y > y_{DM,50})$

$$P_e\left(X > x_{DM,50}\right) = 1 - F_X^e\left(x_{DM,50}\right)$$

$$P_e\left(Y > y_{DM,50}\right) = 1 - F_y^e\left(y_{DM,50}\right)$$

• Based on $P_e(X > x_{DM,50})$ and $P_e(Y > y_{DM,50})$, determine (possibly by interpolation) $P_{e,\rho_{X,Y}^e}\left(X > x_{DM,50} \middle| Y > y_{DM,50}\right)$

Step 5 : Take linear pooling of $P_{e,\rho_{X,Y}^e}\left(X > x_{DM,50} \middle| Y > y_{DM,50}\right)$ to find

$$\pi_{\frac{1}{2},\frac{1}{2}}^{DM}(X,Y) = P\left(F_X^{DM}(X) > \frac{1}{2} \middle| F_Y^{DM}(Y) > \frac{1}{2}\right)$$

Step 6 : Find $\rho_{X,Y}^{DM}$ from the table as the value which corresponds to $\left(\frac{1}{2},\frac{1}{2},\pi_{\frac{1}{2},\frac{1}{2}}^{DM}(X,Y)\right)$.

References

1. B.C.P. Kraan. Probabilistic inversion in Uncertainty Analysis and related topics. Phd Thesis 2001. Delft University of Technology

2. B.C.P. Kraan and R.M. Cooke. Post-processing techniques for the Joint CEC/USNRC Uncertainty Analysis; Journal of Statistical Computation and imulation. 1997. vol. 57 p.243-260.

3. R.M. Cooke. Parameter fitting for uncertain models; Modelling uncertainty in small models. Reliability Engeneering and System Safety. 1994. vol. 44 p-89-102.

4. B.C.P. Kraan and R.M. Cooke. Uncertainty in compartmental models for hazardous material-a case study. Journal of Hazardous Material. 2000. Vol. 71 p. 253-268.

5. A.M.H. Meeuwissen and R.M. Cooke. Tree dependent random variables (revised version). Delft University of Technology. TUD-REPORT 94-28 and ISSN 0922-5641}, 1994. Delft, The Netherlands.

6. J.R. Van Dorp and A. Vinovarov. UNICORN (UNcertainty analysIs with CORrelatioNs); user's manaul.1992. Delft University of Technology.

7. G. Apostolakis and S. Kaplan. Pitfalls in risk calculations. Reliability Engineering & System Safety. 1981. Vol. 2, p.135-145.

8. B.C.P. Kraan and R.M. Cooke. The effect of correlations in uncertainty analysis: two cases. Technical Committee Uncertainty Modelling: Benchmark Workshop. 1997. Editor R.M. Cooke.

9. D. Gokhale and S. Press. Assessment of a prior distribution for the correlation coefficient in a bivariate normal distribution. Journal of the Royal Statistical Society.1982. Vol. 145-2, p.237-249.